LIBRARY
AUG 14 1926

Vol. IX, No. 4 *PSYCHOLOGICAL REVIEW PUBLICATIONS* AUGUST, 1926

Journal of Experimental Psychology

EDITED BY

MADISON BENTLEY, UNIVERSITY OF ILLINOIS
HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
JOHN B. WATSON, NEW YORK CITY (*Review*)
SHEPHERD I. FRANZ, UNIV. OF CALIF., SO. BR. (*Monographs*)
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*Bulletin*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY, Managing Editor

CONTENTS

Studies in Psychometric Theory: ELMER CULLER 271
XIII. *The Normal Law; its Place in Quantitative Psychology*

Specific Serial Learning; A Study of Remote Forward Association: HULSEY CASON 299

An Observational Study of Anger: GEORGINA S. GATES 325

Visual Perception in the White Rat: GLENN D. HIGGINSON 337

The Effect Upon Dark-Adaptation and Visual Periodicity of Atropin and Homatropin: ROLAND C. TRAVIS 348

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

Psychological Review Publications

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
JOHN B. WATSON, 244 MADISON AVE., NEW YORK (*Review*)
SHEPHERD I. FRANZ, UNIV. OF CALIF., SO. BR. (*Monographs*)
MADISON BENTLEY, UNIVERSITY OF ILLINOIS (*J. of Exp. Psych.*)
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*Bulletin*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index*)

WITH THE CO-OPERATION OF

MANY DISTINGUISHED PSYCHOLOGISTS

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bimonthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 480 pages.

PSYCHOLOGICAL BULLETIN

containing abstracts, critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly, the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bimonthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 480 pages.

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in May, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL MONOGRAPHS

consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

Philosophical Monographs: a series of treatises more philosophical in character.
Library of Genetic Science and Philosophy: a series of bound volumes.

ANNUAL SUBSCRIPTION RATES

Review: $5.00 (Foreign, $5.25). **Review and Bulletin:** $9.50 (Foreign, $10.00)
Journal: $5.00 (Foreign, $5.25). **Review and Journal:** $9.00 (Foreign, $9.50)
Bulletin: $5.50 (Foreign, $5.75). **Journal and Bulletin:** $9.50 (Foreign, $10.00)
Any one of above with **Index:** $1.50 additional.
Review, Bulletin, and Journal: $14.00 (Foreign, $14.75).
Review, Bulletin, Journal, and Index: $15.00 (Foreign, $15.75).
Current Numbers: Review or Journal, $1.00; Bulletin, 60c; Index, $2.00
Psychological Monographs: $6.00 per volume (Foreign, $6.30).
Current Issues: prices vary according to size.

Subscriptions, orders, and business communications may be sent direct to the

PRINCETON, N. J. OFFICE

PSYCHOLOGICAL REVIEW COMPANY

FOREIGN AGENTS: G. E. STECHERT & CO., London (2 Star Yard, Cary St., W.C.), Paris (16, rue de Condé)

ExP

Journal of
Experimental Psychology

VOL. IX., No. 4 August, 1926

STUDIES IN PSYCHOMETRIC THEORY[1]

BY ELMER CULLER

University of Illinois

XIII. THE NORMAL LAW; ITS PLACE IN QUANTITATIVE PSYCHOLOGY

In the present critical and experimental study we propose briefly to examine the *empirical* evidence for the normal law in psychometric theory and practise, to sift a few a priori objections to the law, and to formulate as well as we can its true position and import in quantitative psychology.

The normal law implies a theory of natural causation which has been called "beautiful enough to excite worship" (*12*, 43). Take a set of unbiased coins; the proportion of heads in a number of throws will fall in close accord with the terms of the binomial $(\frac{1}{2} + \frac{1}{2})^n$, n being the number of coins in a toss. Now if the statures or Alpha-scores of an unselected group (*1*, 256; *10*, 68) are found to apportion in the same way, we may, by a plausible analogy, think of nature with a handful of unit-factors, all of which contribute to the height or the abilities of an individual and any of which may be present or absent, plus or minus, in a given case. Each member of the species may then be conceived as incorporating the combination of features that turn up in a single toss of nature's hand; in every throw she decides the make-up of one man. The determinants of any physical or mental trait may thus be taken to follow the laws of 'chance' as exhibited in the simple

[1] Of the present series nos. I–X have appeared in (*5*) and nos. XI–XII in this JOURNAL, 1926, 9, 169–194.

and fundamental situation: present or absent. The fascinating simplicity of these conceptions, however, makes them liable to uncritical acceptance and unwise use; so that exacting and judicious appraisals of the normal law and its place in quantitative psychology are always in order.[2]

Everyone knows by this time that empirical distributions in biometry may be far from normal; and theoretical arguments have also been offered to disprove the general applicability of the gaussian hypothesis; for decades it has been decried now and again as a deceptive and exploded fetish of the past. Meanwhile the 'law' is freely accepted and postulated throughout quantitative psychology (particularly in psychophysics and in every kind of scale and test) and its pragmatic value is widely conceded. If the writer's observation be true there seems, in brief, to be a growing divergence of practise from theory. As often happens, men formally question or impugn what in use they continue to cherish. In theory the normal law may be somewhat of a bruised reed; but in practise we still lean heavily upon it. The Pearson formulæ, by which we may graduate any empirical frequencies whatever their degree of asymmetry or kurtosis, are elegant and imposing; but they look too strange and dire for common use and their predictive value is often dubious. Though an occasional writer (Thomson, Pearl) makes earnest with them, most workers in our field leave them carefully alone. More and more the gaussian formula seems to be taken as a 'practical' hypothesis, however debatable its logical adequacy. The psychometric status of the normal law being thus dubious and ill-defined, the evidence herebelow is proposed as an aid in finding its true use and import for quantitative psychology. The argument comprises five heads:

1. The form of any distribution being a function of the measuring-scale to which the frequencies apply, is there any natural or basal *unit of measure* for psychometric purposes (*e.g.*, tests of sensitivity, judgment, intelligence, achievement)? If so, what is it?

[2] Boring gives (*3*) an historical and critical sketch on which some of our comment will focus.

2. Can related measures (*e.g.*, weight, girth, stature) of the same objects all distribute with equal normality?

3. What is the *factual* probability that psychometric data (*e.g.*, differential sensitivity and judgments of relation) will apportion normally?

4. Critical formulation of normal theory, and

5. General survey and conclusions.

1. *What is a natural or 'true' unit of measure in psychometric work?*

What Boring calls the "problem of the unit" (*3*, 20), which he amplifies and reiterates in various forms, is the spearhead of his attack on the normal law. "If we alter the unit of measurement we alter the form of the frequency-distribution . . . it is strange that a relation so obvious should have received so little attention" (p. 18). "Suppose then that nature does conform to the normal law; we shall not in practice be fortunate enough to obtain normal curves of distribution unless we have chanced to use nature's own unit of measurement, or some unit directly proportional to it" (p. 15). He quotes with approval from Bertrand that if a group of measures distribute normally, their logs or squares will not; whence the form of the curve is clearly a function of the measuring-scale. Pearson's statement that the normal law "really does describe in a remarkable manner the distribution of most characters in mankind" and that "for practical purposes we may with nearly absolute safety assume it as a first approximation" (*16*, 142) he protests because we "do not know anything about the units with which we are working" and says that Pearson is "not clear on the part played by the unit of measure" (*3*, p. 23 f). In psychophysics, finally, we can neither generalize nor prophecy because again everything depends on the unit.[3] "Were the psychometric function to approximate the phi-function of gamma ten times, there would be no reason to expect it the eleventh. Suppose the eleventh case were the differential limen for auditory intensity with the Fechner sound pendulum as the instrument. Will the normal law hold? Possibly when we use as the unit $\sin^2(\theta/2)$. If so, it will not hold if we use θ as the unit. If it holds for neither, it may hold for some other function of θ. . . . It is mere luck if the normal law applies" (p. 28).

We find it impossible to share this agitation over the unit. In the auditory case, what are we concerned about? The actual energy or force of the vibrations where they impinge upon and activate the receptive tissue; the work done to the

[3] Boring here reverts to a favored distinction of the Cornell school between "two kinds" of psychophysics. The difference between them may apparently be stated as follows: when we compare two weights by kinesthesis alone or by pressure alone, we are judging "sensations"; when we compare by means of the total impression or configuration, we are judging "stimuli." In article X on the stimulus-error (*5*) we have tried to examine this distinction; and have found the stimulus-attitude no less valid and useful than the other, even though the criteria of judgment are by no means identical in the two procedures.

sensory mechanism. This being hard to gage directly, we are driven to measuring that phase of the physical process on which intensity of vibration depends, *viz.* the pendulum-bob's height of fall.[4] The stimulus ought in theory to be stated in units of force or work (dynes, ergs); in default of these true measures, we resort to one which varies directly with them [height of fall in cm or $\sin^2(\theta/2)$]. Bertrand's comments on the measuring-scale have likewise no appreciable relevance in actual scientific work.[5]

Suppose we have n sets of material, ranging in weight from zero to the capacity of a given scale; these are weighed one by one. Now if the addition of a fixed bit of material, λ, advances the marker, whatever its position on the scale, just one cm in every case, however large or small be the whole mass of material on the platform; then a given interval of the scale (one cm, say) may properly be called a true and natural unit of measure with this instrument. A true scale or unit of measure is always defined by the following condition: equal intervals in every part of the scale stand for equal units of energy or material; in the above weighings *e.g.* the scale-distances correlate perfectly ($r = +1.00$) with the actual units of mass or weight. On a squared or logarithmic scale, this principle fails to hold; in the first, λ displaces the marker more and more as the total weight increases; and in the second less and less. Any rule or standard wherein the same unit of energy or mass is always represented by the same unit of scale in any part of the total range has a fundamental status which holds of no other kind of scale or unit and which thereby warrants the appellation 'true' or 'normal.' Any scale whose relation to the material of measurement is non-linear (*e.g.*, logarithmic) has a derived and secondary place. When using a logarithmic scale we can by computation or with a key find how many real units are involved, by reducing the measures

[4] This equals $1 - \cos\theta = 2\sin^2(\theta/2)$, where θ is the pendulum's deviation from vertical; cf. Fechner (7, 180).

[5] Of this he is himself well aware: "Malgré les objections précédentes, la formule de Gauss doit être adoptée. L'observation la confirme; cela doit suffire dans les applications. Les conséquences minutieusement étudiées se sont toujours trouvées d'accord avec les faits." (2, 178.)

to linear terms; but it is these alone which have ultimate significance. We cannot therefore admit Boring's implication that all units are on the same plane, that our choice of a scale is arbitrary. In speaking of distributions, there is but *one* that has any real significance at all: the frequencies that go with the actual units of energy working upon the receptive tissue or, in default of these, units which correlate perfectly with them. The question is ever this: how do these true units allocate their frequencies? do they show a normal sequence? not how do their logs, roots or other functions behave. In the auditory case the question is, how are the ordinates apportioned that go with abscissæ of $(2 \sin^2 \theta/2)$, since these abscissæ vary in linear relation with energy of incidence at the ear; $\sin \theta$ or any other random function of θ has nothing whatever to do with the matter.

So much for the problem when stimulus and response (the two variables whose relation is defined by the distribution-curve) can be independently measured; but much of the time we have to do with items whose difficulty ('stimulus-value') can be measured only by way of the testee's performance (at what age, or by what proportion of a sampling group, the problem can be solved). We can find the stimulus-value of a lot of weights quite apart from any observer's judgment; but the difficulty of a given problem has no meaning except as revealed in some individual's or group's response. There being no independent scale of stimulus-energy, with which to compare the testee's performance, are we justified in calling the several items of a mental test a real scale, composed of equidistant units, at all? Suppose the true [6] test-value of all the tasks in a given section of army-Alpha (directions, relations) were carefully graded from least to hardest; no. 1 would then differ from 2 by some interval Δ_1, 2 from 3 by another step Δ_2, and so on; save by chance, few of these would be really equal, even though a good many were about the same. In distributing the frequencies, then, by what warrant do we space the scores equally on the base-line, thereby assuming that $210 - 200 = 110 - 100 = 10 - 0$ in difficulty?

[6] That is to say, the mean value for whatever population we are working with.

Here is a *point d'appui* of many critics, with whom Boring seems to be aligned.[7]

Let us treat the question with a simple experiment. Forty numbers (five sets of eight each) ranging from 0 to 10 are set down in random order:

1. 0 — 4 — 9 — 5 — 8 — 3 — 4 — 1
2. 2 — 3 — 7 — 0 — 2 —10 — 6 — 4
3. 4 — 4 — 9 — 2 — 3 — 6 — 5 — 3
4. 6 — 0 —10 — 7 — 2 — 6 — 8 — 4
5. 8 — 5 — 0 — 7 — 1 — 6 — 8 — 2

These figures give the serial differences (Δ) in the lengths of forty lines; the unit of measure is 1/32 inch and all the rows begin with 96 (3 inches). The actual distances are as below:

1. 96 — 100 — 109 — 114 — 122 — 125 — 129 — 130
2. 98 — 101 — 108 — 108 — 110 — 120 — 126 — 130
3. 100 — 104 — 113 — 115 — 118 — 124 — 129 — 132
4. 102 — 102 — 112 — 119 — 121 — 127 — 135 — 139
5. 104 — 109 — 109 — 116 — 117 — 123 — 131 — 133

Their magnitude being thus arbitrarily fixed, the forty lines were carefully ruled with india ink on drawing paper of uniform size (2 x 6.1 inches) backed with cardboard; they were properly centered, being an inch from each side and their midpoints equidistant from the ends of the paper. The cards when shuffled were disposed at random in a well-lighted field, rows of four lengthwise and columns of ten crosswise, each being separated from its neighbor by circa one inch at the end and one-fourth inch on the sides; when not in use, the whole was covered with a screen. The obs. stood directly before the table-surface, where he could move about and survey the whole field with ease. These instructions were read:

> You will be shown forty lines of various lengths, displayed at random within a field and drawn on cards alike in shape and size. You are asked to look with care

[7] "In every case the measure [of intelligence] is arbitrarily chosen; . and it is not even strictly a measure, in the sense of being the sum of equal units, unless we make it so on *a priori* grounds. In an intelligence-test, for example, we have no notion whether the difference between 99 and 100 points is the same difference in intelligence as the difference between 9 and 10 points, unless we are ready to define intelligence arbitrarily as the ability to achieve points in that particular test." (3, 25.)

over the *whole* field, noting the lines in every part; then proceed to arrange them in order of length. Pick up the cards, one by one, choosing always the shortest line still remaining in the field, until all have been taken in hand. You will then have the whole set ranged in order from shortest to longest.

Please remember: (1) to note the lines in every part of the field while making your choice; (2) after choosing a card, not to view it again nor return it to the table. Allcomparisons are to be made in the *field*, not in the hand.

In true order the lines advanced from 96 to 130 by these steps (Δ):

0	1	2	3	4	Total
9	23	6	0	2	40

The mere combination of these five groups, with their random and irregular Δ, reduces the range of Δ from 0–10 to 0–4 and brings a large preponderance of small steps (0 and 1). Upon computing the mean value of the forty Δ and their m.d. from this mean, for the true order as well as for O's rankings, we have:

	Mean Interval (Δ)	M.d.
True Order	1.075	.562
First ranking	1.075	4.275
First two rankings	.988	2.488
First three "	.933	1.847
First four "	.906	1.614
First five "	.950	1.355

With increase of rankings, m.d. steadily and clearly recedes: the trend is so obvious in these five trials that no more rankings were made. Our only concern being to constate whatever tendency or principle proves to be operative, there was no occasion to work through the present problem in further detail. The m.d. will in theory continue to fall pari passu with $1/\sqrt{n}$, and may thus be reduced as far as we please by increasing the number of independent rankings; by the time a thousand subjects, *e.g.*, have made one trial each, the m.d. will be circa 1.355/14.14 or about 0.1; far less than the true m.d. of .562. This means in a word that our forty units or steps, Δ, are becoming more and more nearly equal as the trials proceed.

What does all this mean for our general problem? Let the forty lines be a mental test of forty items, in five groups of eight each; the span of a line standing always for the difficulty

('stimulus-value') of a single test-problem; a testee may be said by his performance to choose the questions that are easiest for him, just as here he proceeds to pick out the shortest lines. We assume furthermore (*a*) that none of the five component groups begin on the same level of difficulty (96, 98, 100, 102, 104); (*b*) that their total range from low to high is not the same (34, 32, 32, 37, 29); (*c*) that the problems advance in difficulty by steps (Δ) ranging from 0 to 10; and (*d*) that these Δ are ordered in a random and arbitrary way. The problem of equality in the units (Δ) is by our conditions posed in a flagrant form; it is impossible to believe that the steps (Δ), from 1 to *n*, in any well made test can vary so widely and arbitrarily as here.[8] By purposely exaggerating the problem we make sure that our conclusions hold a fortiori in all ordinary tests. Now we see, upon merely combining the five groups, that the range of Δ reduces by 60 per cent. (from 0–10 to 0–4) and the true intervals begin to approximate equality; and that furthermore as they are worked over in successive trials, the m.d. consistently falls and can be reduced as far as we please by multiplying the number of testees. In a mental test, where the items can be spaced with a fair degree of approximation to begin with, the several steps 1 – 2, ——, 101 – 102, ——, 201 – 202, can then be equalized with whatever degree of precision we need. If the initial Δ be not grossly unequal, their eventual equality for psychometric purposes can be attained with relative ease.[9]

[8] Because the difficulty of the several tasks may be measured in advance of the final scale with a refinement limited only by the care and resources of the test-maker.

[9] The operation of Weber's law in the testing field has never been adequately examined; the oft quoted dictum ("differences that are equally often noticed are equal, unless the differences are either always or never noticed;" *cf. 8*, 18) is acceptable only with grave reservations; were this rule applied in ordinary differential perception (weights, temperatures), we should be misled into concluding that 2 (j.n.d. at 100) equals 4 (j.n.d. at 200) equals 6 (j.n.d. at 300) and so on. In the case of a mental test, suppose a pair of the harder problems are just perceptibily unlike in difficulty and likewise a pair of the easier; or suppose two good samples of English composition are just noticeably unequal in merit and likewise two poor ones; is the 'limen' or discriminable difference larger in the pair that lies farther away from zero? We cannot say with assurance; but the writer ventures that Δ *does* grow pari passu with a growth in difficulty or in merit. If this be true, the test- or scale-units, instead of remaining equal from zero upward, enlarge by a constant ratio; in a scale like

The question of the present section may then be answered in this way: (*a*) when dealing with material that is 'objectively' measureable (weights, temperatures), the 'true' measure, with which alone we are concerned in psychometry, is one which correlates perfectly (varies in linear relation) with the actual amount of energy that is working upon a given receptive tissue of the organism; (*b*) with material whose objective value is unknown (tests, scales of merit), the scale-units can be equalized with whatever precision our needs require and resources permit. When so equalized, these units may be treated as 'true' measures in the sense of (*a*),[10] since their stimulus-value is taken directly from its effect on the organism's behavior.[11]

The business of psychometry, then, is to find the relation between these units or measures of *stimulus* and the correlative units of *response*;[12] their connection or co-variation is described by the distribution-curve. If we hold firmly to this conception, many of our difficulties vanish.[13]

Hillegas', *e.g.*, for measuring quality in English composition, the units, instead of being equal "in the same sense, though not as accurately, as millimeters" (*8*, 13), would be many times as large at the upper limit (937) as near zero.

The whole question of Weber's law in this field calls for review and search; but a systematic factor of this kind would be very unlike the problem we faced above, and hence need not delay us here.

[10] In practice of course they are neither so precise nor easily attained.

[11] 'Stimulus' can be defined only in terms of an organism which in some way is modified by or reacts to it. There is accordingly no more fundamental way of finding stimulus-value than via the organism's response.

[12] Viewed from its effect on the organism, a stimulus varies in four dimensions: kind (visual, auditory), degree (amount of work done within unit-area), extent (amount of receptive tissue affected), duration (period through which it continues to act). This means, the reader will note, that a stimulus may be measured and stated in units of *time* (seconds), provided only the other dimensions remain constant. The baseline of a test or scale may thus be composed of time-units. This fact, being sometimes forgotten, is emphasized here.

The term 'response' is here broad enough to include any modification of the subject's experience or activity in consequence of the stimulation.

[13] *E.g.*, our units of stimulation are not "arbitrary" (*3*, 31); they are defined as above and can be replaced only by other units which correlate perfectly with them. Neither have we any concern with a so-called "psychological unit" (p. 31) except as a phase or stage of the response. The more fully and minutely the organism's reaction is known, to be sure, the more fundamental does our work become; but the *stimulus* is always the primary variable.

2. *Can related measures (e.g., weight, girth, stature) taken from the same class of objects all distribute with equal normality?*

Suppose we have a lot of cubical crystals of common salt, and measure the weight and one edge of each sample. Since the weight should vary as the cube of the edge, "it is plain that the normal law could not hold in both cases." If nature seeks to conform to the normal law, does she prefer to conform in lengths or in volumes or in some other unit? "If man would seek the law of nature, he must have cogent reason for his choice of measures" (3, 16). Having heard the same contention from more than one accepted statistician in psychology the writer begs to offer a little evidence on the matter.

Suppose we have a thousand cubes whose edges are normally apportioned; how may their volumes be expected to distribute? If they are conceptual (mathematical) cubes, where by definition $V = x^3$ (x being the edge), then of course the volumes are not possibly normal; they will skew negatively, trailing off to the low end and bunching at the upper. Here we are concerned however not with mathematical fictions but with facts of nature; in science as well as statecraft, it is always "a condition, not a theory" that meets us. In each case we have to do with curves described by a set of *empirical* values; and it is readily conceivable that both volumes and edges may apportion normally; for the reason that the stated relation $V = x^3$ does not precisely hold with actual cubes; their volumes always deviate from x^3 by a margin of error. The type of argument above noted makes no adequate account of the fact that no two natural objects of the same class (cubes, human bodies) have exactly the same shape; they are not like conceptual figures where one dimension immutably fixes all the rest. In a mathematical cube, given the edge, you know also the volume; but with any existing cube on earth, given one edge the volume is not precisely known.

How does the above analysis accord with experimental evidence? We may examine two cases. (1) Consider Pearl's data on the physical characters of the egg, wherein the length,

breadth and weight of 450 eggs were precisely measured. If the principle above quoted be true, weight may be expected to vary as the cube of *both* length and width; so that if weight distributes normally the two linear dimensions will both skew in the same direction, whereas if these are both normal the weight will skew to the same side of each. Now a normal curve, the reader will note, is a transitional type in the Pearson family and has two necessary and sufficient indices: symmetry ($\chi = 0$) and mesokurtosis ($\eta = 0$). Deviations from symmetry are measured by Pearson's formula for skewness, χ (cf. *12*, 277) and the degree of kurtosis (peakedness) by the value of β_2 (in a normal curve $\beta_2 = 3$, in all other curves $\beta_2 \neq 3$). Pearl's table of analytical constants (*13*, 184) shows at once that in both criteria, skewness and kurtosis, the figure for weight (W) falls between those for length (L) and breadth (B):

	L	W	B	PE
Skewness	0.2155	0.1704	0.0718	0.0389
Kurtosis ($\beta_2 - 3$)	1.3732	0.2448	−0.1168	0.1556

W is skewed less than L and more than B; it is more platykurtic (flat-topped) than L and less so than B. Were these relations mathematically determined W, being proportional to both L^3 and B^3, would fall above or below both alike, in degree of skewness and kurtosis. This being untrue, it follows that the above principle is in no sense crucial here. Both L and W, we may note, are significantly skew and L likewise leptokurtic.

We may stalk the problem from another angle, by finding the general goodness of fit (Pearson's P); whereupon we have:[14]

[14] The reader unfamiliar with curve-fitting may consult Pearson (*17*, xxxi) or Rietz (*19*, 78), and should note that P gives but an approximate, though highly useful, value. It is approximate for some three reasons: (*a*) while Pearson's procedure for curve-fitting is standard, it can hardly be called the sole definitive method; others may well be possible which would give somewhat variant results (*cf. 19*, 110); (*b*) it is subject, like any statistical magnitude, to errors of random sampling; (*c*) its size

	L	*W*	*B*
χ^2	17.02	9.60	11.82
No. of cells	11	10	12
P	.07	.39	.38

B and *W* may properly be called 'good' fits while *L* is pretty poor though still tolerable.[15] Two magnitudes *B* and *W*, which are supposed to relate as x to x^3, both turn out to be equally normal; we are freed of the choice which Boring would impose of trying in advance to divine which of the two units, cubic or linear, nature prefers to distribute normally; either will serve as nature's measure equally well. Not all characters will of course distribute with equal normality; but in case weight is normally apportioned and length is not or vice versa, we must *look* for the reason; it cannot be ascribed offhand to some cubic or other supposed functional relation that would make it impossible for both to distribute normally at the same

depends upon the way in which the tails of the surface are treated. The principle here is to avoid extremely small adjusted values else we are likely to get abnormally large or small values for χ^2 and *P*. Suppose, *e.g.*, a theoretical ordinate is 0.1 and the observed frequency 1 (a plausible situation, since if there is any frequency at all it cannot be less than one unit); they differ by 0.9 which, when squared and divided by 0.1, gives 8.1—a figure so large it might easily reduce *P* from say .50 to .10, from a good to a poor fit. The actual discrepancy between observed and theoretical frequency (0.9) is quite negligible, but the effect on *P* is tremendous because the denominator is so minute. It becomes then a matter of judgment how many cases from the tails shall be combined so as to avoid the diminutive divisor and at the same time not cover up any important tail-characters of the curve. Our procedure we believe is conservative; the ordinates of each tail are combined, counting inward from the extremes, until they equal not less than 2 per cent. of the total frequency. So in the curve for breadth, *N* being 450, 2 per cent. = 9. At the low end of the distribution (*13*, 183; Table 4) we find in order 2, 4, 9, 18, ..., the correlative theoretical frequencies being 2.0, 4.6, 10.1, 18.7, ... By combining the first three figures (2.0 + 4.6 + 10.1 = 16.7 as vs. the observed frequency 15), we get not less than 9. The 2 per cent.-rule insures a sizable divisor and, by reducing the number of cells in the whole distribution, makes the value of *P* all the more conservative and dependable (Pearson commonly combines about 1 per cent. in each tail; *cf.* *17*, xxxii). In fitting $\Phi(\gamma)$ percentile curves by Thomson's method (*20*), we treat every theoretical percentage as the nearest whole number; *e.g.*, 89.4 equals 89.0.

[15] "*P* equals .39" (in the case of *W*) means: if we took 100 samples at random from a distribution known to be exactly normal, 39 of them would depart from normality at least as much as the present curve for *W*. This makes it quite possible that *W* is 'really' normal.

time. In the present case then this principle does not operate in any decisive way.

(2) Consider another case; certain data from the anthropometric committee of the British Association (*1*, 285) on length and weight of infants at birth. The measures apparently proceed from staff medical attendants in two hospitals and would seem to be fairly reliable, weight being taken without clothing and stature in a recumbent position. Here are the values of P for length and weight:

	Males		Females	
	L	*W*	*L*	*W*
No. of subjects	451		466	
χ^2	1.94	3.09	10.77	4.55
No. of cells	6	6	6	6
P	.86	.69	.06	.48

Two of the fits are indeed excellent, one is distinctly good and one is pretty poor; and we also find that with males length is more normal, with females weight. Neither one nor other can then be called "nature's own unit"; one distributes in the mean about as normally as the other; and despite their rough cubic relation both may fit the gaussian law extremely well.

From all this we by no means conclude that every distribution is gaussian; far from it; but anormality if it appears cannot be reduced offhand to a non-linear relation with something else. Every curve must be taken for itself; if one surface be normal and another not, we must seek out the reason and not try to find it in some conceptual relation to another variable. Nowhere in nature do we find anything like an exact theoretical correspondence of characters such as length and volume; it is quite possible then for two imperfectly correlated variables (weight and stature) both to distribute normally.

3. *What is the empirical likelihood of a normal curve in psychometric work?*

We are ready now for the real question: How do various kinds of psychometric data actually distribute? Is the gaussian merely one of an infinite variety of curves, being no more probable nor typical than any of the rest? No informed person nowaday believes that all distributions in nature are even roundly normal; as well contend that "all men are created equal" or that "the sun shines every day in London." What we need, and here aim to provide, is a little empirical evidence on the way frequencies actually fall within the field of psychophysics [16] and inferentially the wider field of quantitative psychology as a whole. Our concern is purely psychometric: how do the responses of a human organism distribute in a situation that can be defined and controlled within our limits of knowledge, and from which we are careful to banish any factor that may disturb our frequencies; may the normal law with reasonable confidence be assumed in quantitative work where the responses proceed from a known and stabilized milieu instead of the unknown and haphazard conditions often imposed by nature?

Turn to the evidence. In article XI (*6*) of this series are shown the values of P (degree of approximation to the gaussian formula) in four sets of independent data from seven observers with two types of material (temperature and lifted weights) and varied procedures; they were given as a fair sample of what may be expected in differential work. We now examine these and a fifth case to find how many good,

[16] So-called psychophysical data (differential sensitivity and perception of relations) are ideal for this end, because:

(*a*) the experimental milieu is more simple and stable than are the complex and variable conditions of mental testing;

(*b*) stimulus and response (abscissa and ordinate) are measured in complete independence of each other. With many scales, on the contrary [*e.g.*, Hillegas (*8*) and Thorndike (*21*)], a normal distribution of the judges' ratings is first assumed and the unit of measure then derived by way of that postulate. These units are clearly unfitted to our need.

(*c*) the units of stimulation are known with great exactitude. Although the steps in a mental test or achievement scale can also be equalized and measured (*cf.* above), yet this can never be done of course with a precision and ease comparable to finding weights and temperatures.

average and poor fits (P) appear in this representative body of data.

We first need to inquire: What value must P have to rate as a 'good' fit, such that a normal curve may, for all scientific purposes, be used to graduate and describe the empirical frequencies? From an urn containing 50 white and 50 black balls, we draw say a thousand sets of n each recording every time the proportion of whites. Even though the urn is expressly filled so as to yield normal frequencies $(\frac{1}{2} + \frac{1}{2})^n$, we should never in practise have a perfect fit; all would deviate much or little from the norm because of sampling errors. Now the equation, $P = .80$, means: if we drew 100 samples (each comprising 1,000 sets of n balls) from this urn which 'ought' to give a truly gaussian set of frequencies, 80 of them would depart from the norm by as much as (or more than) the curve in question. All would agree that, when $P = .80$, theory and fact are in splendid accord. But many cases are not so clear; is P still good at .60 or .40 or .20? The meaning of 'good' and 'poor' fit is clearly a matter of evaluation and judgment, which can be answered only from the usage of accepted authorities. The examples tabulated below and gleaned from men whose competence will not be questioned, may serve to show how these experts characterize particular values of P.

Pearson

$P = .89$	"excellent" ..... (*17*, xxxiii)
.80	"excellent" ..... (*18*, 365)
.64	"excellent" ..... (*17*, xxxii)
.56	"good" ..... (*18*, 365)
	"remarkably good" ..... (*14*, 170)
.45	"excellent" ..... (*18*, 367)
.37, .69	"In the face of such probabilities as this any sound statistician would not hesitate to say" that the data "obey the normal law of frequency" ..... (*15*, 343)
.28	A normal curve "fairly represents" the distribution ..... (*14*, 174)
.21	"quite fair" ..... (*18*, 367)
.18	"quite fair" ..... (*18*, 367)
.02	"bad" ..... (*17*, xxxii)

Thomson

.91	"excellent" ..... (*4*, 91)
.48	"good" ..... (*4*, 91)
.007	It is "not at all probable" that the fitted curve truly represents the actual distribution ..... (*20*, 226)

Rietz

.31 "good" ..... (*19*, 81)
.10 "not bad" ..... (*19*, 80)

Pearl

.65 "very fair, and certainly entirely satisfactory for all practical purposes" ..... (*11*, 42)

Without multiplying citations we venture to propose the following table of characterizations for *P*:

1.00–.75	Superlative
.74–.50	Excellent
.49–.25	Good
.24–.10	Medium or fair
.09–.05	Poor
Below .05	Unacceptable

P becomes unacceptable at .05; a degree of deviation from normality that would occur five times in a hundred even were we sampling a situation known to be truly gaussian (drawing from a 50–50 urn). This measure of a-normality further is just as likely to appear as a deviation of 2.45 p.e. from the mean; a probability of 3.0 p.e., as conservative workers often demand, would lower the figure to about .02. The index of acceptability is therefore generously high; and the whole table in our opinion will command assent from those who are qualified to judge.

Let us now turn to the data; the following tabulations may be verified from the basic figures in Art. XI (*6*).

I. CULLER: 92 distributions in all, 46 warmer and 46 cooler

	W	*C*	Both	Percentage	
				Individual	Cumulative
Superlative	19	17	36	.391	.391
Excellent	9	3	12	.131	.522
Good	4	12	16	.174	.696
Fair	3	6	9	.098	.794
Poor	2	2	4	.043	.837
Unacceptable	9	6	15	.163	1.000

It appears that 70 per cent. of the whole number of distributions, covering all stages of practise at all temperatures and all space-time orders with both forms of response *W* and *C*, were good fits and over half were excellent; this is all the more significant because of those which are unacceptable over half (8 in 15) come at 20°; for some reason unknown to the writer the ordinates at this level are abnormally erratic, particularly in the second series; the figures for *P* are in the mean far below those on either side (16° and 24°).

II. URBAN'S THREE OBSERVERS: 18 distributions for each, 54 in all

	I	II	III	Total	Percentage	
					Individual	Cumulative
Superlative	1	7	6	14	.259	.259
Excellent	4	1	4	9	.167	.426
Good	2	6	2	10	.185	.611
Fair	5	4	2	11	.204	.815
Poor	3	0	1	4	.074	.889
Unacceptable	3	0	3	6	.111	1.000

Here 61 per cent. of the whole number for all Os with various degrees of skill and training and using both types of response (heavier and lighter) are good or better. The most proficient in the group (II, Urban himself) has none below fair, while over three-fourths (77 per cent.) are good. III furthermore was improving in both proficiency and fit as practise

III. FERNBERGER'S OBSERVERS II AND VIII: 56 curves for each, 112 for both

	II	VIII	Both	Percentage	
				Individual	Cumulative
Superlative	20	30	50	.446	.446
Excellent	16	7	23	.205	.651
Good	13	13	26	.232	.883
Fair	2	2	4	.036	.919
Poor	2	2	4	.036	.955
Unacceptable	3	2	5	.045	1.000

proceeded, so that his mean degree of normality was notably higher at the close of the nine series than at the beginning (Art. XI).

Here 88 per cent. of the whole number is good or better: additional proof how well a competent observer approximates normality in his differential judgments.

IV. Brown's Observer: 18 curves tested

Of these, 9 are superlative, 5 excellent, 4 good; 100 per cent. being therefore good or better. These cases were chosen by a fixed rule (Art. XI) so as to reveal practise improvement; hence they represent varying stages of proficiency.

V. In his acoumetric study Keller (*9*, 49) used Wundt's fallphonometer, the standard stimulus (*S*) ranging from 40 to 60 cm (height of fall of sphere upon ebony plate), the comparison-stimuli going from 12 to 24 cm above and below *S* by 3-cm steps. For present needs the two extreme modes of response, deutlich kleiner (*DK*) and deutlich grösser (*DG*), are more useful than the rest; partly because they provide a form of judgment not represented in the four sets of data above. From his tables we lift only the percentages of *DK* and *DG* for four observers (Ba, Me, Se, We) and three values of *S* (45, 50, 55 cm); these alone are extensive enough for our needs and, being the most complete, may likewise be taken as the most truly representative of Keller's data. Precisely the same figures are used by Urban as well in his methodological study; from his tables (*23*, 374–379) we can easily find the values of *P*, to-wit:

S =	45		50		55	
	DK	*DG*	*DK*	*DG*	*DK*	*DG*
Ba	.88	.79	.30	.69	.92	.79
Me	.19	.41	.69	.03	.91	.16
Se	.96	.36	.95	.32	.50	.74
We	.35	.69	.56	.67	.56	.34

These 24 cases divide as follows:

	Ba	Me	Se	We	Total	Percentage	
						Differential	Cumulative
Superlative	4	1	2	0	7	.292	.292
Excellent	1	1	2	4	8	.333	.625
Good	1	1	2	2	6	.250	.875
Fair		2			2	.083	.958
Poor						.000	.958
Unacceptable		1			1	.042	1.000

For all observers over 87 per cent. are good or better; which is the more noteworthy since all below good are due to one obs. (Me); he seems to be less trained than the others (*cf.* p. 52). In any event the figures agree with what has been shown before.

The above material comprises just 300 empirical distributions apportioned as below:

	Total	Empirical Percentages		Theoretical (50–50 urn) Percentages	
		Differential	Cumulative	Differential	Cumulative
Superlative	116	.387	.387	.250	.250
Excellent	57	.190	.577	.250	.500
Good	62	.207	.783	.250	.750
Fair	26	.087	.870	.150	.900
Poor	12	.040	.910	.050	.950
Unacceptable	27	.090	1.000	.050	1.000

These figures have more to say than many words. Had we drawn the reports (greater and not-greater) from a universe purposely designed to give normal frequencies (a 50–50 urn) instead of from human beings, P would have exceeded .75 but one time in four (25.0 per cent.) and .50 but one time in two; here on the contrary the proportions rise to 38.7 per cent and 57.7 per cent. Only one-half of the urn-curves would do better than $P = .50$, but 73 per cent. (8 in 11) of these observers do so. These figures, be it remembered, are drawn from five independent sources: eleven observers, three types of material (temperatures, sounds, lifted weights), two kinds of reports ('greater' and 'distinctly greater'), unlike procedures; a true

and fair sample, we take it, of the field. In face of these reports on differential sensitivity and perception of relation (whether we call them judgments of 'stimulus' or judgments of 'sensation'), it is hard to believe that normality is a matter of "luck" (*3*, 28).[17] We are, on the contrary, enabled to say that a careful observer, working with stated instruction and uniform attitude in a situation experimentally controlled, tends so insistently toward normality that *we are more likely to find a normal sequence in his reports than when drawing balls from a 50–50 urn.* Here we conclude the argument; merely staying to remind the reader that if normality is low to begin with it tends to improve with time (Art. XI, *6*); as practise proceeds and proficiency mounts, degree of normality rises notably as well.

4. *Theoretical formulation*

The conventional urn is beyond doubt a simple and useful means of concretizing questions of probability. With an urn-schema the basal condition of normality is 50 per cent. white and 50 per cent. black balls; if this be true in every trial, the proportion of *W* (or *B*) will describe a gaussian surface provided n be large and the samples many $(\frac{1}{2} + \frac{1}{2})^n$. It may well seem, however, that the 50–50 proportion is no more likely than any other; and that in general the composition of a given urn is wholly arbitrary, fixed by the personal *Willkür* of some experimenter and therefore with no value for natural phenomena and their causation. It behooves us then to look for a moment at the key-position of the 50–50 relation in the arch of statistical theory, not alone but in any system of natural forces as well.

Consider a closed field of energy; the whole system being self-contained, no energy is imparted or withdrawn ab extra. Being thus free of interference from without, action and re-

[17] Boring is careful to say that normality appears by "luck" only with introspective reports on mental process and not with judgments of stimulus; but it seems we can seldom be sure to which class a body of data belongs. Boring is not sure about Urban's material (*3*, 28) and the writer is equally unsure how his own data would be characterized by a 'stimulus-error' expert; so that we cannot venture to say how many of the above cases are properly 'introspective' and how many are improperly 'objective.'

action within the field are equal and opposed in every direction (space-continuum)and at every moment (time-continuum); the energy-vectors in one part of the universe at any time will be compensated by others of unlike sign elsewhere. The integral energy of the system may for simplicity be conceived as centering in and acting from the centroid or point of equilibrum (γ), all the lines of force radiating from γ being represented by vectors of suitable direction and extent. Now let the system contract without limit, remaining always impervious to energy from without; as the field closes in upon γ, the sum-total of these energy-vectors within the space-time continuum will approach a sphere of perfectly uniform energy-density, centering in γ. This limiting condition is of course ideal; the lines of force will as a rule not be equal in every direction from the centroid. But in any event the whole energy-content may be divided into equal and opposed halves by any plane which cuts the system through γ; the forces impinging upon this plane from one side (+) will precisely balance those acting upon it from the other (−); were this untrue, displacement of γ would follow. The positive and negative units of energy are equal in number, just as are the black and white balls in a 50–50 urn; our closed field may therefore be treated as an ideal urn $(\frac{1}{2} + \frac{1}{2})^n$, with the properties thereto pertaining.

This 50–50 relation holds, be it noted, only for the centroid, where goes on a *perfectly stable* interplay of forces; it need not (is indeed very unlikely to) be true for most other parts of the field, where the flow of energy may take many forms (*e.g.*, circular about γ). It being in practice impossible to find the energy-vectors in every part of the system and equally hopeless to locate the centroid, we are fain to sample the moments of force at random points and for various times, in the hope that these samples may fairly represent the whole. We test the space-time continuum as well as our needs seem to require and resources to permit; and were the samples truly representative of the *whole* field we should *always* get a normal surface just as when we draw from an ideal 50–50 urn; but in practice we may with our limited samples get any kind of frequency-

curves. At the point we use for testing, there may be a positive flow of energy set off by a negative flow elsewhere; we should then draw a preponderance of plus units (*W* balls), and our frequencies would be skewed from the low end of the scale. Or the points we sample may happen to give an excess of positive units near .50 and few at .40 and .60, thus making our curve leptokurtic, too peaked in the mode and too low on either side. Every kind of distribution is clearly possible with restricted samples from a large system; and yet in every case the empirical frequencies would depart from normality only because they give us no fair picture of the whole field whence they are drawn and of which they are presumed to be representative.

When, by reason of limited sampling, a distribution turns out to be parakurtic or asymmetrical, we are not warranted in fitting it à la Procrustes to a normal pattern; we should resort to Pearson's generalized formula. The value of the Pearson-curves is twofold: (*a*) practical, by enabling us to graduate any body of material and so learn what to expect in future samples from the same limited and instable population; (*b*) theoretical, by measuring the degree and direction wherein our empirical curve departs from the norm. Only the latter concerns us here. The normal curve is indeed a *norm* just as the Venus de Milo, *e.g.*, in artistic tradition; which by no means implies that all frequencies are gaussian nor yet that all female figures display the classic lines of the Melian goddess; but a norm has the great value of showing what is 'wrong' with a case which fails to conform. So the Pearson-constants, by revealing how and in what degree a curve departs from normality, show where the problems of causation lie. A normal surface, being just what may be expected of a stable system with its balanced interplay of forces, sets no problem in distribution; whereas anormality is a mark and proof that some special agent or trend is operative in the field we have sampled. Knowing that the stature of Scottish insane, male and female, is leptokurtic (*22*, 309–311), which means that a disproportionate ratio are of medial height while those above and below the mid-range are too few, we at once face the

problem: have stature and insanity a causal relation? does medial height favor the incidence of mental defect? From the graph alone we cannot be sure; the same leptokurtosis may characterize the whole group whence these defectives are drawn. If the underlying population prove on investigation to be normal, then we know that medial stature *does* take a role in the etiology of Scotch insanity; whereas if the general population be also leptokurtic, our problem is merely rephrased: have stature and 'Scottishness' a causal relation? what is there in the Scotch nature or milieu to favor medial height? In any event the cardinal fact about the present stature-curve is its anormality; whence we know without more ado that some factor (which can only be defined by further research) is causing the insane statures to deviate from what is commonly found in the population at large. Wherever skewness and kurtosis appear, we may *ohne weiteres* infer that some agency is interfering with the balanced interplay of action and reaction in an ideal system; the deviation in a word becomes a clue or key for localizing the forces at work.

The normal law then, with its auxiliary Pearson-constants, provides an easy and extremely useful norm for finding how well a given situation (experimental or natural) accords with a stable or ideal system of forces in mutual and balanced interaction; while the degree and direction of an empirical curve's departure from normality show where its problems of causation lie.

5. *General survey and conclusions*

Let us review the way we have come.

1. Within a stable and self-contained system, action and reaction are equal and opposed; the whole may therefore be regarded as an ideal urn with 50 per cent. white and 50 per cent. black balls, or as a binomial $(\frac{1}{2} + \frac{1}{2})^n$.

2. A truly representative sampling of this field of energy will always give a normal set of frequencies, even as does the binomial and the ideal urn.

3. The frequencies we draw will be normal only if equal units of scale in the sample stand for equal units of energy in

the field; a perfect correlation of energy-content with units of measure is always presupposed; with roots, logs and other non-linear functions we have nothing to do. In drawing balls from a 50–50 urn, n stands for the actual number of balls in a single draw, not for a power, root or log of that number.

4. If variable A (stature) ensues from one complex of forces and a correlated variable B (weight) from another, there is no evidence that normality of A will affect the distribution of B in any particular sense; from the A-curve nothing can be foretold about the frequencies of B. Both may be equally gaussian, or one may excel the other; but the partial and imperfect correlation between them (B varies roughly with A^3) is no explanation. Every complex of conditions must be taken for itself alone, much as the constitution of any urn can be learned only by adequate sampling; when anormality appears, the cause must be found by empirical research; it cannot be imputed to partial and crude relations with something else.

5. When a distribution turns out to be anormal, we know ipso facto that *its frequencies proceed from an unstable set of conditions*, from a milieu which is out of equilibrium; that the stable interplay of causal factors, which characterizes an ideal system, does not hold of the stimulus-complex whence these responses flow. The measures of anormality (Pearson) show whether and in what degree our experimental or natural setting deviates from the ideal condition of equilibrium, and how its instability modifies the data. Any departure from the norm becomes a problem in causation; since, when a complex of forces is more and more screened from interference ab extra, its expressions tend also toward the norm.

6. In Article XI we saw that degree of normality rises notably as practise proceeds and proficiency mounts; and here we have tried to show that, in one of the most important fields of psychometry (differential sensitivity and perception of relation), where the stimulus-values are independently known, the reports of a qualified Obs. are more likely to show a normal sequence than do the draws from a 50–50 urn. In

this empirical evidence the conclusions of theory are sustained. In quantitative psychology we try to achieve a closed field of conditions, whence are excluded with great care all variable factors that are known and amenable to our control; of our Obs. and his milieu we try to make a unified system in stable equilibrium: to be sure that he have no preference for one response (heavier) over the other (lighter), that his degree of proficiency (practise) remain virtually constant during the trials, that he be free of changes in general well-being, in attitude and interest, in expectation or knowledge of results, and the like. Insofar as a stable setting is thereby achieved we may rightfully even though tentatively postulate a normal distribution; whereas, if anormal frequencies continue to appear, the writer would not scruple to infer without hesitation that an important variable is still 'at large.' Prior to the event, of course, we can never know that our Obs. (field of energy) has been duly shielded from all the variables that may disturb his responses. But this is a restriction that the future always imposes; even in science, doubt and error take toll of prediction. But we know at least what a normal curve in theory implies; and when we have imposed all known conditions for stabilizing an experimental setting, we are warranted, as a "first approximation" (*16*, 143), in assuming normality of response.

7. To psychometry (and biometry) the principles herein set forth would seem to apply more neatly and fully than to any other field of science; for the reason that a human being (or any organism) is an integrated system of energy, with a finely balanced interplay of action and reaction; and it is further the activity or responses of the *integral* organism that makes our raw material; whence we see that normality may well be expected as a feature of human function. Behaviorism has been differentiated from physiology in the proposition that the latter has to do with specific organs, the former with activity wherein the whole being participates (*24*, 20); and the Gestalt-theorie insists even more on the integral nature of experience and action. If this be true, an organism's activities eventuate from, and thus typify or sample, its whole field

of energy instead of a part or phase; and by our criterion should therefore evince normality of response.

8. In a given case, one of the Pearson-curves will commonly meet the empirical frequencies more nearly than does the gaussian formula; and some have therefore urged that we graduate our experimental data by the more elaborate Pearson-methods, and thus get for every body of material a unique formula. If our interests are practical it may be of use to do so, and thus find what may be expected in future samples of the same anormal population; just as an insurance company might wish to predict the future mortality rate in a given 'anormal' district; but for purposes of scientific theory, this proposal has little merit. In all scientific generalization we are commonly concerned, not with finding a formula that will exactly accommodate a few sampling observations, but rather a curve which will more roughly (and yet adequately to our needs) describe a great many. We can always find a formula to fit every one of our empirical frequencies; but such a curve, being without predictive value, is sterile. Now its keenest opponent would hardly contend that any other member of the Pearson-family is so widely useful as the gaussian. Psychometric curves, when they depart from normality, show little consistency: one is skewed to right, another to left, one is too peaked, another too flat; each has its own peculiarities, but they commonly cluster all about the symmetrical and mesokurtic normal limit. Indeed we get a surprising ratio of good approximations to normality, even in experimental situations that are poorly controlled. If we make up with reasonable care a mental test and try it on a random sample, our frequencies are virtually certain to be more like a gaussian than like any other equally general surface. In a purely heuristic way then it seems the most fruitful to use.

We see again that the Pearson-methods are of use, not so much to graduate and in a sense 'fix' the anormality as to disclose whether and in what degree the latter exists, thereby apprizing us that our field is still imperfectly stabilized and open to variable irruptions. Every beginner in physics works

out the law of a falling body, and none, save by chance, ever gets precisely $V = \frac{1}{2} gt^2$; but nobody suggests that these diverse values be accepted as founding a law different from the standard or 'normal' law of fall; instead we say that air-resistance and mechanical defects and inaccurate readings have adulterated the data. The physicist is not concerned about the thousand rates of fall that one gets with diverse barometric readings and instruments and observers; rather is he concerned with the limiting value which appears more nearly as conditions are standardized. In like manner, scientific psychometry is primarily concerned with the limiting case about which many empirical curves cluster; and is secondarily concerned with the direction and degree wherein a particular curve departs from this limiting norm.

To conclude: we have tried hereabove to show

(*a*) that in quantitative psychology (as opposed to natural phenomena at large) we are very likely to obtain frequencies which can be represented with all needed precision by the normal formula, and

(*b*) that in case the frequencies prove to be anormal, this formula (supplemented with the Pearson-constants) remains equally valuable for revealing the mode or kind of anormality, finding whether it is statistically significant, and thereby apprizing us whether and in what degree our experimental situation is inadequately stabilized and controlled.

Quantitative psychology has, as we believe, no single concept more widely useful than the normal law, when correctly understood and applied.

REFERENCES

1. ANTHROPOMETRIC COMMITTEE. Report of the British Association for the Advancement of Science, London, 1883.
2. BERTRAND, J. Calcul des probabilités. (2^{me} éd.) Paris, 1907.
3. BORING, E. G. Logic of the normal law of error in mental measurement. *Amer. J. of Psychol.*, 1920, **31**, 1.
4. BROWN, W., & THOMSON, G. H. Essentials of mental measurement. Cambridge, 1921.
5. CULLER, ELMER. Studies in psychometric theory, I–X. *Psychol. Monog.*, 1926, **35**, no. 2, 56–137.
6. CULLER, ELMER. Studies in psychometric theory, XI–XII. This JOURNAL, 1926, **9**, 169–194.

7. FECHNER, G. T. Elemente der Psychophysik, Band I. Leipzig, 1907.
8. HILLEGAS, M. B. Scale for the measurement of quality in English Composition by young people. *Teachers Coll. Rec.*, 1912, **13**, no. 4.
9. KELLER, H. Die Methode der mehrfachen Fälle im Gebiete der Schallempfindungen, usw. *Psychol. Stud.*, 1907, **3**, 49.
10. OTIS, A. S. Statistical method in educational measurement. Yonkers, 1925.
11. PEARL, R. Variation and correlation in brain-weight. *Biometrika*, 1905, **4**, 13.
12. PEARL, R. Medical biometry and statistics. Philadelphia, 1923.
13. PEARL, R., & SURFACE, F. M. Variation and correlation in the physical characters of the egg. *Bull. U. S. Bureau of Animal Industry*, 1914, **110**, part III.
14. PEARSON, K. On the criterion that a given system of deviations, etc. *Phil. Mag.*, 1900, **50** (ser. 5), 157.
15. PEARSON, K. Craniological notes. *Biometrika*, 1903, **2**, 339.
16. PEARSON, K. On the inheritance of the mental and moral characters in man, etc. *Biometrika*, 1904, **3**, 131.
17. PEARSON, K. Tables for statisticians and biometricians. Cambridge, 1914.
18. PEARSON, K., & LEE, A. Inheritance of physical characters. *Biometrika*, 1903, **2**, 357.
19. RIETZ, H. L. Handbook of mathematical statistics. Boston, 1924.
20. THOMSON, G. H. Criterion of goodness of fit of psychophysical curves. *Biometrika*, 1919. **12**, 216.
21. THORNDIKE, E. L. Handwriting. *Teachers Coll. Rec.*, 1910, **11**, no. 2.
22. TOCHER, J. F. Anthropometric characteristics of the inmates of asylums in Scotland. *Biometrika*, 1907, **5**, 298.
23. URBAN, F. M. Über die Methode der mehrfachen Fälle. *Arch. f. d. ges. Psychol.*, 1910, **17**, 367.
24. WATSON, J. B. Psychology from the standpoint of a behaviorist. (2d ed.) Philadelphia, 1924.

SPECIFIC SERIAL LEARNING; A STUDY OF REMOTE FORWARD ASSOCIATION

BY HULSEY CASON

University of Rochester

Preliminary Statement

The theory of remote association assumes that, when a series is learned so that the items can be reproduced in the proper order, certain remote associations are formed between each item and all other items, in both the forward and backward directions. Ebbinghaus claimed that the remote associations are stronger in the forward than in the backward direction, and that the strength of the associations decreases with the number of intervening items.[1] We have submitted this problem to an experimental investigation, and the results are summarized in two reports. The first [2] is concerned with the problem of backward association, while the present article describes experiments in which the problem of remote forward association is for the first time submitted to a systematic experimental investigation.[3] The two papers should be considered together, since they are concerned with different aspects of the same problem of remote association.[4]

Part I. Tests with Difficult Prose Material

Introduction

If the series *1-2-3-4-5-6-7-8-9-10* is so learned that the items can be reproduced in the proper order, does this learning result in the formation of remote forward associations and facilitate the later learning of the items in the order *1-3-5-7-9*,

[1] See literature cited on page 195 *supra* (this Journal, June, 1926).

[2] See this Journal, 1926, 9, 195–227.

[3] I wish to make a grateful acknowledgment to my wife, Eloise Boeker Cason, for much valuable assistance in the preparation of the manuscript, and to the 33 subjects for their coöperation in a very tiresome procedure.

[4] We are not concerned in the present studies with *mediate association*.

or *2-4-6-8-10, etc.?* Where the subject actually practises the items in the order *1-3-5-7-9*, while the principal task of learning the series in the order *1-2-3-4-5-6-7-8-9-10* is in progress, any associations formed in the order *1-3-5-7-9* are of course "direct" or "immediate," rather than "indirect" or "remote." The theory of remote association assumes that when associations are formed in the order *1-2-3-4-5-6-7-8-9-10*, associations are *at the same time* established in the order *1-3-5-7-9*.

In the present experiment the subjects were required to learn difficult prose material, and appropriate tests were given to determine whether remote forward associations had also been formed.

Material

The prose passage selected is an adaptation from Dewey's "Human nature and conduct" (pp. 106–107). It is quoted in full as follows:

SOME WRITERS HAVE MADE THE IDEA OF THE COMPLETE MALLEABILITY OF HUMAN NATURE THE BASIS FOR ASSERTING THE OMNIPOTENCE OF EDUCATION TO SHAPE HUMAN SOCIETY. OTHERS HAVE FOUND IN THESE PLANS FOR SOCIAL CHANGE EVIDENCES OF THE PRONENESS OF YOUTH TO ILLUSION, OR OF INCAPACITY ON THE PART OF THOSE WHO HAVE GROWN OLD TO LEARN ANYTHING FROM EXPERIENCE. CURIOUSLY ENOUGH BOTH PARTIES REST THEIR CASE UPON JUST THE FACTOR WHICH WHEN IT IS ANALYZED WEAKENS THEIR RESPECTIVE CONCLUSIONS. THAT IS TO SAY, THE RADICAL REFORMER RESTS HIS CONTENTION IN BEHALF OF EASY AND RAPID CHANGE UPON THE PSYCHOLOGY OF HABITS, OF INSTITUTIONS IN SHAPING RAW NATURE, AND THE CONSERVATIVE GROUNDS HIS COUNTER-ASSERTION UPON THE PSYCHOLOGY OF INSTINCTS. AS MATTER OF FACT, IT IS PRECISELY CUSTOM WHICH HAS GREATEST INERTIA, WHICH IS LEAST SUSCEPTIBLE OF ALTERATION; WHILE INSTINCTS ARE MOST READILY MODIFIABLE THROUGH USE, MOST SUBJECT TO EDUCATIVE DIRECTION. THE CONSERVATIVE WHO BEGS SCIENTIFIC SUPPORT FROM THE PSYCHOLOGY OF INSTINCTS IS THE VICTIM OF AN OUTGROWN PSYCHOLOGY WHICH DERIVED ITS NOTION OF INSTINCT FROM AN EXAGGERATION OF THE FIXITY AND CERTAINTY OF THE OPERATION OF INSTINCTS AMONG THE LOWER ANIMALS. HE IS A VICTIM OF A POPULAR ZOÖLOGY OF THE BIRD, BEE AND BEAVER. HE IS IGNORANT THAT INSTINCTS IN THE ANIMALS ARE LESS INFALLIBLE AND DEFINITE THAN IS SUPPOSED, AND ALSO THAT THE HUMAN BEING DIFFERS FROM THE LOWER ANIMALS IN PRECISELY THE FACT THAT HIS NATIVE ACTIVITIES LACK THE COMPLEX READY-MADE ORGANIZATION OF THE ANIMALS' ORIGINAL ABILITIES.

This passage was selected because it is difficult to learn, and also because, at the time the experiment was carried out, most of the 17 subjects were taking the writer's course in social psychology, in which the question of biological and social heredity had just been discussed.

For the method which was used in deriving 11 forward lists and 11 separated lists from the prose selection the reader is referred to the similar procedure for backward lists on page 214. The first 6 and the last 3 words of the prose selection were omitted, leaving 242 words. The first forward list was made up of the 11 alternate words in the F_1 (forward) row, and each of the remaining *F*-lists, F_2, F_3, *etc.*, was constructed in the same way. The first separated list was made up of the 11 words in the column above "S_1." Each of the *F*- and *S*-lists was typewritten on a separate card.

Procedure

The writer acted as experimenter, and the subjects were taken individually. The procedure with each subject may be divided into 3 parts: (1) Learning the *F*- and *S*-lists, (2) Learning the prose passage, and (3) Relearning the *F*- and *S*-lists. Comparisons were made between the efficiency of learning these lists, both before and after learning the prose passage.

Learning the F- and S-Lists.—After learning 3 practise lists, the subjects learned the 11 *F*-lists and the 11 *S*-lists in the order F_1, S_1, F_2, S_2, *etc.* He was given one list at a time, and was instructed to study it aloud by the whole method until he could make a correct oral recitation. A record was made of the time required to learn each list. The conditions of learning were kept as uniform as possible, and an effort was made to control the situation in such a way that the subjects would not know that the lists were made up of words taken from a logical prose passage and arranged as described above. No subject noticed the method used in constructing the lists.

Learning the Prose Passage.—A printed copy of the prose passage was given the subject, and he was instructed to learn it in his own way. After the prose passage had been learned

to the point of one correct recitation, the subject was required to study it further and to recite it daily for a period of one week. The passage proved to be very difficult; but in learning it the subject formed very strong associations between the consecutive words or groups of words. It would seem at first thought that the distributed practise on this long difficult passage should afford ample opportunity for the formation of remote forward associations.

Relearning the F- and S-Lists.—After the prose passage had been relearned and recited daily for a period of one week, the subject relearned these lists, using the same method as before.

After completing, each subject was asked a number of questions about the experiment. All thought that most of the words in the 25 lists had been taken from the prose passage, but they did not notice the method used in constructing them, and no subject had any insight into the problem being investigated. All subjects stated that various parts of the prose passage would intrude themselves, or were recalled from time to time, while they were relearning the lists. This tendency was more pronounced in some lists than in others. Almost all of the subjects thought that this recall of parts of the prose passage facilitated the relearning of the lists, by making them more familiar, and giving them more meaning. The objective results, however, showed that these opinions were not reliable.

The *S*-lists were made up of words arranged in the *backward* direction, when the words in these lists are compared with the arrangement of the same words in the prose passage. They are composed of words which were originally separated from each other in the prose passage by *21* intervening words. The *F*-lists, on the other hand, are made up of words arranged in the *forward* direction, and separated from each other by only *1* intervening word, when they are compared with the order of the words in the prose passage. Now when the prose passage is learned, it can be assumed that the principle of remote association should establish much stronger associations between every 2d word in the *forward* direction (*F-lists*) than between every 22d word in the *backward* direction (*S-lists*).

It can be assumed in fact that *the strength of the associations between every 22d word in the backward direction (S-lists) is practically zero.*

The S-lists represent a *zero base of learning* in the present experiment, and the presence or absence of remote forward associations may therefore be determined on the basis of a comparison between the efficiency of learning the S-lists and the *F*-lists, before and after learning the prose passage. If the efficiency of learning the *F*-lists (compared with the S-lists) *improves* after learning the prose passage, a conclusion in favor of remote forward association is justified. If learning the prose passage *reduces* the efficiency of learning the *F*-lists (compared with the S-lists), then it follows that learning the prose passage inhibits the formation of remote forward associations.

Results

Influence of Learning Difficult Prose Passage on Remote Forward Associations; Individual Results.—The individual results for the 17 subjects are summarized in Tables I*A*, *B* and *C*. The general method of treating the data may be illustrated by the results of the first subject (Table I*A*).

The first column in the table shows that before learning the prose passage the 11 *F*-lists required an average learning time of about 86 sec, and the 11 S-lists an average learning time of about 97 sec. The superiority of the *F*-lists over the S-lists therefore is about 11 sec. This difference is expressed in per cent. by dividing 11 by the average of 86 and 97.

After learning the prose passage, the same *F*-lists required an average learning time of 55 sec, and the same S-lists an average of 63 sec. The difference is now + 8 sec, in favor of the *F*-lists.

If the figures in the last two lines of the table are positive, the conclusion is in favor of remote forward association. If these figures are negative, the results indicate that learning the prose passage has resulted in a condition which inhibits the formation of remote forward associations. The percentage difference given in the last row of the table is more significant than the difference $(d–c) - (b–a)$, because of the shift in the base of comparison after learning the prose passage.

TABLE I*A* AND I*B*

INFLUENCE OF LEARNING DIFFICULT PROSE PASSAGE ON REMOTE FORWARD ASSOCIATIONS; INDIVIDUAL RESULTS

Subjects	1		2		3		4		5		6	
Central Tendency	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.
Before Learning Passage:												
(*a*) *F*-Lists	86.3	81.0	53.9	48.0	83.0	72.0	88.4	83.0	57.4	47.0	50.9	46.0
(*b*) *S*-Lists	97.7	84.0	61.5	41.0	100.6	72.0	77.6	70.0	65.2	59.0	54.1	51.0
b–*a*	+11.4	+ 3.0	+ 7.6	− 7.0	+17.6	0.0	−10.8	−13.0	+ 7.8	+12.0	+ 3.2	+ 5.0
(Per cent.)	+12.4	+ 3.6	+13.2	−15.7	+19.2	0.0	−13.0	−17.0	+12.7	+22.6	+ 6.1	+10.3
After Learning Passage:												
(*c*) *F*-Lists	55.2	47.0	34.2	30.0	40.6	38.0	59.3	57.0	44.5	42.0	28.2	26.0
(*d*) *S*-Lists	62.9	48.0	40.1	35.0	63.2	45.0	65.5	48.0	44.5	46.0	32.0	36.0
d–*c*	+ 7.7	+ 1.0	+ 5.9	+ 5.0	+22.6	+ 7.0	+ 6.2	− 9.0	0.0	+ 4.0	+ 3.8	+10.0
(Per cent.)	+13.0	+ 2.0	+15.9	+15.4	+43.5	+16.9	+ 9.9	+17.1	0.0	+ 9.1	+12.6	+32.3
Effect of Learning Passage, or (*d*–*c*) − (*b*–*a*)	− 3.7	− 2.0	− 1.7	+12.0	+ 5.0	+ 7.0	+17.0	+ 4.0	− 7.8	− 8.0	+ 0.6	+ 5.0
(Per cent.)	+ 0.6	− 1.5	+ 2.7	+31.1	+24.3	+16.9	+22.9	+34.1	−12.7	−13.5	+ 6.5	+22.0

Subjects	7		8		9		10		11		12	
Central Tendency	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.
Before Learning Passage:												
(*a*) *F*-Lists	65.2	58.0	71.6	68.0	124.5	125.0	100.0	70.0	81.3	75.0	57.7	58.0
(*b*) *S*-Lists	64.8	65.0	91.0	81.0	107.2	86.0	103.2	99.0	94.3	87.0	75.9	71.0
b–*a*	− 0.4	+ 7.0	+19.4	+13.0	−17.3	−39.0	+ 3.2	+29.0	+13.0	+12.0	+18.2	+13.0
(Per cent.)	− 0.6	+11.4	+23.9	+17.4	−14.9	−37.0	+ 3.1	+34.3	+14.8	+14.8	+27.2	+20.2
After Learning Passage:												
(*c*) *F*-Lists	33.7	29.0	41.6	43.0	54.0	45.0	71.8	67.0	53.8	57.0	38.5	28.0
(*d*) *S*-Lists	38.7	43.0	45.1	40.0	63.7	62.0	81.2	76.0	58.5	49.0	50.2	42.0
d–*c*	+ 5.0	+14.0	+ 3.5	− 3.0	+ 9.7	+17.0	+ 9.4	+ 9.0	+ 4.7	− 8.0	+11.7	+14.0
(Per cent.)	+13.8	+38.9	+ 8.1	− 7.2	+16.5	+31.8	+12.3	+12.6	+ 8.4	−15.1	+26.4	+40.0
Effect of Learning Passage, or (*d*–*c*) − (*b*–*a*)	+ 5.4	+ 7.0	−15.9	−16.0	+27.0	+56.0	+ 6.2	−20.0	− 8.3	−20.0	− 6.5	+ 1.0
(Per cent.)	+14.4	+27.5	−15.8	−24.6	+31.4	+68.8	+ 9.2	−21.7	− 6.4	−29.9	− 0.8	+19.8

TABLE IC

INFLUENCE OF LEARNING DIFFICULT PROSE PASSAGE ON REMOTE FORWARD ASSOCIATIONS; INDIVIDUAL RESULTS

Subjects	13		14		15		16		17	
Central Tendency	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.
Before Learning Passage:										
(*a*) *F*-Lists	75.1	44.0	50.3	41.0	60.5	57.0	95.6	94.0	37.7	37.0
(*b*) *S*-Lists	104.1	107.0	105.2	69.0	52.4	48.0	104.6	89.0	57.5	51.0
b–*a*	+29.0	+63.0	+54.9	+28.0	− 8.1	− 9.0	+ 9.0	− 5.0	+19.8	+14.0
(Per cent.)	+32.4	+83.4	+70.6	+50.9	−14.3	−17.1	+ 9.0	− 5.5	+41.6	+31.8
After Learning Passage:										
(*c*) *F*-Lists	37.4	41.0	36.0	30.0	42.1	35.0	61.0	54.0	32.0	24.0
(*d*) *S*-Lists	47.3	44.0	43.4	26.0	39.9	48.0	77.3	67.0	34.9	30.0
d–*c*	+ 9.9	+ 3.0	+ 7.4	− 4.0	− 2.2	+13.0	+16.3	+13.0	+ 2.9	+ 6.0
(Per cent.)	+23.4	+ 7.1	+18.6	−14.3	− 5.4	+31.3	+23.6	+21.5	+ 8.7	+22.2
Effect of Learning Passage, or (*d*–*c*) − (*b*–*a*)	−19.1	−60.0	−47.5	−32.0	+ 5.9	+22.0	+ 7.3	+18.0	−16.9	− 8.0
(Per cent.)	− 9.0	−76.3	−52.0	−65.2	+ 8.9	+48.4	+14.6	+27.0	−32.9	− 9.6

These tables show that there is considerable variability in the results of the individual subjects. Some give results which are definitely in favor of remote forward association, while the results of others are just as definitely opposed to this theory. If the values for $(d–c) – (b–a)$ are considered, half of the subjects give results in favor of remote forward association, and half give the opposite results. If the percentage differences are considered, the results are very slightly in favor of the theory.

There is very little correlation between the type of the results and the nature of the individual subjects. Thus the most naive subjects are nos. 3, 6, 7, 8, 10, 13, 16, and 17, 4 of whom give results for and 4 results against the principle of remote forward association. The subjects best versed in psychology are nos. 9 and 14. No. 9 gives results in favor, and no. 14 results against the theory. The most conscientious subjects are nos. 6, 7, 13, and 17. Two of these give results for, and 2 against. Five of the subjects could be classified as slightly "peculiar" (nos. 1, 5, 11, 15, and 16). Two of these give results for, and 3 against. The most distinctly "abnormal" subject in the group is no. 15, who gives results in favor of the theory. The uncontrolled factors in the situation and the use of different learning methods seem to afford a plausible explanation for the variability of the results.

Influence of Learning Difficult Prose Passage on Remote Forward Associations; Combined Results of 17 Subjects.—All of the gross time scores for the 17 subjects are arbitrarily combined in Table II. The general method of presentation is already familiar. The table shows that if the values for $(d–c) – (b–a)$ are considered, the results are slightly opposed to remote forward association. If the percentage differences are considered, the results are very slightly in favor of the theory. The results as a whole seem to indicate that, with the procedure used, the chances of obtaining a result in either direction are practically equal.

Although the variability of the individual results described above is significant from a practical point of view, the results themselves do not exclude the possibility that there may be a

TABLE II

INFLUENCE OF LEARNING DIFFICULT PROSE PASSAGE ON REMOTE FORWARD ASSOCIATIONS; COMBINED RESULTS OF 17 SUBJECTS

Statistical Measure	Ave.	Q_1	Q_2	Q_3
Before Learning Passage:				
(*a*) *F*-Lists	72.9±2.0	43.0	63.5±2.5	94.0
No. cases	187	187	187	187
S. D.	39.7			
(*b*) *S*-Lists	83.3±2.9	50.5	70.0±3.6	104.0
No. cases	187	187	187	187
S. D.	58.0			
b–*a*	+10.4±3.5	+ 7.5	+ 6.5±4.4	+10.0
(Per cent.)	+13.3	+16.0	+ 9.7	+10.1
After Learning Passage:				
(*c*) *F*-Lists	44.9±1.2	27.0	40.7±1.5	57.3
No. cases	187	187	187	187
S. D.	23.5			
(*d*) *S*-Lists	52.2±1.6	32.3	45.2±2.0	64.1
No. cases	187	187	187	187
S. D.	32.7			
d–*c*	+ 7.3±2.0	+ 5.1	+ 4.5±2.5	+ 6.8
(Per cent.)	+15.0	+17.1	+10.5	+11.2
Effect of Learning Passage, or (*d*–*c*) − (*b*–*a*)	− 3.1±4.0	− 2.4	− 2.0±5.0	− 3.2
(Per cent.)	+ 1.7	+ 1.1	+ 0.8	+ 1.1

facilitation or an inhibition of remote forward associations as a result of learning the prose material. This problem is investigated by a different procedure in Part II.

PART II. TESTS WITH EASY LOGICAL MATERIAL

Introduction

An easy prose passage and a poem were used as learning material in a further study of remote forward association. An effort was made to secure greater uniformity in the results and to avoid the individual irregularities which occurred in Part I. The poem used was Robert Burns' "Afton Water," the last verse being omitted. The easy prose passage was an adaptation from the latter part of Robert Louis Stevenson's essay, "Aes Triplex." It is quoted in full as follows:

IT IS BETTER TO LIVE AND BE DONE WITH IT, THAN TO DIE DAILY IN THE SICKROOM. BY ALL MEANS BEGIN YOUR FOLIO; EVEN IF THE DOCTOR DOES NOT GIVE YOU A YEAR, EVEN IF HE HESITATES ABOUT A MONTH, MAKE ONE BRAVE PUSH AND SEE WHAT CAN BE ACCOMPLISHED IN A WEEK. IT IS NOT ONLY IN FINISHED UNDER-

TAKINGS THAT WE OUGHT TO HONOUR USEFUL LABOUR. A SPIRIT GOES OUT OF THE MAN WHO MEANS EXECUTION, WHICH OUTLIVES THE MOST UNTIMELY ENDING. ALL WHO HAVE MEANT GOOD WORK WITH THEIR WHOLE HEARTS, HAVE DONE GOOD WORK, ALTHOUGH THEY MAY DIE BEFORE THEY HAVE THE TIME TO SIGN IT. EVERY HEART THAT HAS BEAT STRONG AND CHEERFULLY HAS LEFT A HOPEFUL IMPULSE BEHIND IT IN THE WORLD, AND BETTERED THE TRADITION OF MANKIND. AND EVEN IF DEATH CATCH PEOPLE, LAYING OUT VAST PROJECTS, AND PLANNING MONSTROUS FOUNDATIONS, FLUSHED WITH HOPE, THEY SHOULD BE AT ONCE TRIPPED UP AND SILENCED: IS THERE NOT SOMETHING BRAVE AND SPIRITED IN SUCH A TERMINATION?

The method used in constructing the forward and separated lists was the same as that described in Part I. Three sets of material were used:

(*a*) The prose passage as given above. 9 *F*-lists and 9 *S*-lists (9 words each) were derived from the first 162 words.

(*b*) The prose passage, with the last sentence omitted. 8 *F*-lists and 8 *S*-lists (8 words each) were derived from the first 128 words.

(*c*) The poem referred to above. 9 *F*-lists and 9 *S*-lists (9 words each) were derived from the first 162 words.

Procedure

The writer acted as experimenter, and the subjects were taken individually in a single experimental period which lasted between 3 and 4 hrs. The procedure used with each set of material consisted of 3 steps: (1) Learning and relearning the *F*- and *S*-lists, (2) Learning the easy logical selection, and (3) Relearning the *F*- and *S*-lists. As in the experiments of Part I, comparisons were made between the efficiency of learning the *F*-lists and the *S*-lists, before and after learning the logical selection.

Learning and Relearning the F- and S-Lists.—The subject learned the *F*- and *S*-lists in the order S_1, F_1, S_2, F_2, *etc.* He was given one list at a time, and was instructed to study it in his own way until he could make a correct oral recitation. A record was kept of the time required to learn each list.

After the lists had been learned as described above, they were relearned 1 or 2 times. In this way we obtained a more

reliable measure of the natural relative difficulty of learning the *F*- and the *S*-lists, before learning the logical selection.

Learning the Logical Selection.—The subject was instructed to study the logical selection in his own way until he could recite it perfectly and without hesitation. When the passage was practically learned, the writer helped the subject in further practise until the recitation was very smooth. The subject practised both silently and aloud.

Relearning the F- and S-Lists.—The subject was finally required to relearn these lists 2 or 3 times, using the same method as before.

When the third and last part of the procedure had been completed, the subject was asked a number of questions about the experiment. None of the subjects noticed the way the lists had been constructed; and no subject had any insight into the problem being investigated. About half of the subjects claimed that parts of the easy logical material would intrude themselves, or were recalled from time to time, while they were relearning the *F*- and *S*-lists in the third part of the procedure.

Before considering the results in the following section, it is desirable to note that the *S*-lists are made up of every *16*th (or every *18*th) word arranged in the *backward* direction, when the words in these lists are compared with the arrangement of the same words in the logical selection. The *F-lists*, however, are made up of every *2*d word arranged in the *forward* direction. Since learning the logical material does not facilitate the learning of the *S*-lists, the presence or absence of remote forward associations can be determined on the basis of a comparison between the efficiency of learning the two kinds, before and after learning the logical material. According to the theory of remote association, the efficiency of learning the *F*-lists, compared with the efficiency of learning the *S*-lists, should *improve* after learning the logical material.

Results

Influence of Learning Easy Prose Passage on Remote Forward Associations; Individual Results.—The individual

TABLE III

INFLUENCE OF LEARNING EASY PROSE PASSAGE ON REMOTE FORWARD ASSOCIATIONS; INDIVIDUAL RESULTS

Subjects	1		2		3		4		5		6		7		8	
Central Tendency	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.
Before Learning Passage:																
(*a*) *F*-Lists	9.87	6.90	11.00	7.00	11.28	8.00	18.78	14.50	13.50	9.50	28.11	21.50	12.04	9.00	18.33	13.00
(*b*) *S*-Lists	10.37	6.70	16.37	11.00	13.03	7.50	19.34	17.50	15.00	8.50	33.06	24.00	14.89	12.00	31.78	23.00
b–*a*	+ 0.50	−0.20	+ 5.37	+ 4.00	+ 1.75	−0.50	+ 0.56	+ 3.00	+ 1.50	− 1.00	+ 4.95	+ 2.50	+ 2.85	+ 3.00	+13.45	+10.00
(Per cent.)	+ 4.94	−2.94	+39.24	+44.44	+14.39	−6.45	+ 2.94	+18.75	+10.53	−11.11	+16.18	+10.99	+21.18	+28.57	+53.68	+55.55
After Learning Passage:																
(*c*) *F*-Lists	5.37	4.21	4.96	4.33	5.12	4.21	8.21	5.10	9.12	6.00	13.17	9.50	6.22	5.43	5.92	4.96
(*d*) *S*-Lists	4.81	4.06	5.75	4.39	6.97	4.00	7.35	5.19	8.96	7.50	13.06	10.25	7.69	7.00	7.58	5.17
d–*c*	− 0.56	−0.15	+ 0.79	+ 0.06	+ 1.85	−0.21	− 0.86	+ 0.09	− 0.16	+ 1.50	− 0.11	+ 0.75	+ 1.47	+ 1.57	+ 1.66	+ 0.21
(Per cent.)	−11.00	−3.63	+14.75	+ 1.38	+30.60	−5.12	−11.05	+ 1.75	− 1.77	+22.22	− 0.84	+ 7.59	+21.14	+25.26	+24.59	+ 4.15
Effect of Learning Passage, or (*d*–*c*) − (*b*–*a*)	− 1.06	+0.05	− 4.58	− 3.94	+ 0.10	+0.29	− 1.42	− 2.91	− 1.66	+ 2.50	− 5.06	− 1.75	− 0.38	− 1.43	−11.79	− 9.79
(Per cent.)	−15.94	−0.69	−24.49	−43.06	+16.21	+1.33	−13.99	−17.00	−12.30	+33.33	−17.02	− 3.40	− 0.04	− 3.31	−29.09	−51.40
Nature of Results	?		−		+		−		+		−		+		−	

TABLE IV

INFLUENCE OF LEARNING POEM ON REMOTE FORWARD ASSOCIATIONS; INDIVIDUAL RESULTS

Subjects	9		10		11		12		13		14		15		16	
Central Tendency	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.	Ave.	Med.
Before Learning Poem:																
(*a*) *F*-Lists	15.30	10.12	22.94	16.00	20.94	16.00	31.78	22.25	16.28	12.50	17.72	16.50	30.06	21.50	41.17	29.50
(*b*) *S*-Lists	16.11	10.87	26.00	22.50	24.33	20.00	29.89	22.00	20.50	12.50	15.61	15.00	34.39	31.50	44.14	30.00
b–*a*	+ 0.81	+ 0.75	+ 3.06	+ 6.50	+ 3.39	+ 4.00	− 1.89	− 0.25	+ 4.22	0.00	− 2.11	− 1.50	+ 4.33	+10.00	+ 2.97	+ 0.50
(Per cent.)	+ 5.16	+ 7.15	+12.51	+33.77	+14.98	+22.22	− 6.13	− 1.13	+22.95	0.00	−12.66	− 9.52	+13.44	+37.74	+ 6.96	+ 1.68
After Learning Poem:																
(*c*) *F*-Lists	6.85	5.00	11.37	7.00	6.44	5.20	11.50	8.50	8.00	6.41	7.41	6.25	16.63	14.00	26.78	24.00
(*d*) *S*-Lists	6.33	5.00	11.00	7.50	5.22	4.80	10.61	8.50	7.59	6.00	7.89	7.00	18.26	13.50	18.56	16.00
d–*c*	− 0.52	0.00	− 0.37	+ 0.50	− 1.22	− 0.40	− 0.89	0.00	− 0.41	− 0.41	+ 0.48	+ 0.75	+ 1.63	− 0.50	− 8.22	− 8.00
(Per cent.)	− 7.89	0.00	− 3.31	+ 6.90	−20.93	− 8.00	− 8.05	0.00	− 5.26	− 6.61	+ 6.27	+11.32	+ 9.34	− 3.64	−36.26	−40.00
Effect of Learning Poem, or (*d*–*c*) − (*b*–*a*)	− 1.33	− 0.75	− 3.43	− 6.00	− 4.61	− 4.40	+ 1.00	+ 0.25	− 4.63	− 0.41	+ 2.59	+ 2.25	− 2.70	−10.50	−11.19	− 8.50
(Per cent.)	−13.05	− 7.15	−15.82	−26.87	−35.91	−30.22	− 1.92	+ 1.13	−28.21	− 6.61	+18.93	+20.84	− 4.10	−41.38	−43.22	−41.68
Nature of Results	−		−		−		?		−		+		−		−	

results of 8 subjects are shown in Table III. The first 5 subjects used the set of material made up of lists of 8 words each; and the last 3 subjects used the longer set of material composed of lists of 9 words each. The general method of treating the data is already familiar, but the changes in procedure require a different interpretation of some of the results.

The average results for subject no. 2 may be considered for purposes of illustration. Before this subject learned the prose passage, the *F*-lists were 39 per cent more efficient than the *S*. After learning the passage, the *F*-lists were 15 per cent. more efficient than the *S*. Learning the passage apparently *reduced* the efficiency of learning the *F*-lists (compared with the *S*-lists) about 24 per cent. This is probably a spurious result, however, because this subject relearned the *F*- and *S*-lists 3 times both before and after learning the passage.

Now relearning the *F*- and *S*-lists several times after learning the passage tends to bring distributions (*c*) and (*d*) closer together; and this reduces the value for *d–c*. If the value for *b–a* is positive and large, the value for (*d–c*) – (*b–a*) will probably be negative; and in this case there may be an erroneous impression that learning the passage tends to inhibit the formation of remote forward associations. If *b–a* is very small, the value for (*d–c*) – (*b–a*) will indicate the true nature of the results,—'plus' meaning for and 'minus' meaning against the principle of remote forward association. If *b–a* is negative and *d–c* is positive, the results are clearly *in favor of* the theory. If *b–a* is positive and *d–c* is negative, the results are clearly *against* the theory.

With these considerations in mind, it is possible to interpret the results of Table III with reasonable accuracy. Subject no. 1 is *against* the theory if the average results are considered, but *in favor of* the theory if the median results are considered. Hence the question mark in the lower row of the table. Subject no. 2 seems to be slightly *against* the theory, considering the small values for *d–c*. No. 3 is *for* the theory because of the large value for *d–c*. The interpretation which we have placed on each subject's results is indicated in the

lower row of the table. Whether the results of the individuals will be for or against the theory seems to depend upon factors which are not controlled and upon the various learning methods used by the several subjects.

Influence of Learning Poem on Remote Forward Associations; Individual Results.—The individual results of 8 subjects are shown in Table IV. The results for no. 14 are in favor of remote forward association; the results for no. 12 are questionable; and the results for the 6 remaining subjects indicate that learning poetry inhibits the formation of remote forward associations.

This inhibition is apparently due to the recall of certain parts of the poem while the subjects are relearning the forward lists in the third part of the procedure. Considering the way the lists are constructed, it seems plausible to believe that the *F*-lists would suggest the poem more often than the *S*-lists. This recall of parts of the poem probably serves as a distraction, and increases the time required to learn the *F*-lists.

As one passes from the results of Part I to the results shown in Tables III and IV, it might be expected that the evidence in favor of remote forward association would increase, since the poetry was easier to learn than the prose passage, and the subjects learned it much more thoroughly; and also because both kinds of material in the present experiment were learned more satisfactorily than the difficult prose passage in Part I. *Instead of an increase, however, we find a decrease in the evidence for remote forward association. When the material is learned very smoothly, as in the case of the poetry, the formation of remote forward associations is inhibited.*

There is some correspondence between the nature of the 16 subjects and the nature of the results which they give. Four of the subjects are graduate students in psychology; nos. 1 (?), 5 (+), 10(−), and 13 (−). No 5 (+) is a foreign student whose knowledge of English is limited. The most intelligent subjects are probably nos. 1 (?), 3 (+), 7 (+), 10 (−), 11 (−), and 13 (−). Subject no. 11 (−) is a boy 13 years of age, in the Junior class in College, and with an I.Q. of about 180. The less intelligent subjects are probably nos.

8 (−), 12 (?), 15 (−), and 16 (−), especially the last-named. The most patient subjects are nos. 2 (−), 6 (−), 9 (−), 10 (−), and 13 (−). The most tense or impatient subjects are nos. 3 (+), 4 (−), 7 (+), 11 (−), 14 (+), and 15 (−). These data suggest that those individuals who are very intelligent, and those who are somewhat impatient, tend to give results in favor of remote forward association; whereas the less intelligent and more patient subjects tend to show an inhibition of remote forward associations.

Influence of Learning Easy Logical Material on Remote Forward Associations; Combined Results of 16 Subjects.—In Table V we have arbitrarily combined all of the gross time scores of the 16 subjects. The combined results of only the 8 subjects who learned poetry would give a very reliable result *against* the principle of remote association; but combining the gross time scores obtained with both types of material gives a result which is perhaps more typical of everyday learning situations, even if the reliability of the results is not so high as it would have been with poetry alone.

TABLE V

INFLUENCE OF LEARNING EASY LOGICAL MATERIAL ON REMOTE FORWARD ASSOCIATIONS; COMBINED RESULTS OF 16 SUBJECTS

Statistical Measure	Ave.	Q_1	Q_2	Q_3
Before Learning Selection:				
(*a*) *F*-Lists	19.20±0.62	6.96	12.66±0.78	24.50
No. cases	436	436	436	436
S. D.	19.26			
(*b*) *S*-Lists	21.81±0.72	7.57	14.42±0.90	27.25
No. cases	436	436	436	436
S. D.	22.20			
b–*a*	+ 2.61±0.95	+ 0.61	+ 1.76±1.19	+ 2.75
(Per cent.)	+12.73	+ 8.40	+13.00	+10.63
After Learning Selection:				
(*c*) *F*-Lists	9.31±0.26	4.45	6.02±0.33	10.38
No. cases	489	489	489	489
S. D.	8.60			
(*d*) *S*-Lists	9.10±0.23	4.55	6.38±0.29	11.33
No. cases	489	489	489	489
S. D.	7.39			
d–*c*	− 0.21±0.35	+ 0.10	+ 0.36±0.44	+ 0.95
(Per cent.)	− 2.28	+ 2.22	+ 5.81	+ 8.75
Effect of Learning Selection, or $(d-c)-(b-a)$	− 2.82±1.01	− 0.51	− 1.40±1.27	− 1.80
(Per cent.)	−15.01	− 6.18	− 7.19	− 1.88

Table V shows that before learning the logical selection the *F*-lists required an average of 19.20 sec, as against 21.81 sec for the *S*-lists. The *F*-lists were therefore *easier* than the *S*-lists, *b–a* being equal to + 2.61 ± 0.95. After learning the selection, the *F*-lists were *more difficult* than the *S*-lists, *d–c* being equal to − 0.21 ± 0.35. This represents a shift in per cent. from + 12.73 to − 2.28. It may also be noted that before learning the logical selection, the *F*-lists were *less variable* than the *S*-lists, but after learning the selection the *F*-lists were *more variable.*

The value for (*d–c*) − (*b–a*), or − 2.82 ± 1.01, supports a conclusion against the principle of remote forward association. The median scores also support the same conclusion; but the reliability of these results is not as high as that of the average results. Considering the large number of measures, and the importance of the extreme deviations in the present study, the average seems to have more logical significance.

Summary

In Part I of the present paper 17 subjects were required to learn difficult prose material, and tests were given to determine whether remote forward associations had also been formed. The practise on the prose passage was distributed, and very strong associations were formed between the consecutive words or groups of words. The tests with the forward lists and separated lists showed considerable variability in the individual results, but the combined results were on the whole opposed to the principle of remote forward association.

In Part II an easy prose passage and a poem were used in a further study of the same problem. The 16 subjects were taken individually in an experimental period which lasted between 3 and 4 hours. The tests with the forward and separated lists were given immediately after the logical material had been learned, and any remote associations which had been formed had the definite advantage of recency. The conditions of the experiment were similar to the learning situations of everyday life inasmuch as the subjects were allowed to use their own methods of study. When the prose

passage was used as learning material, the results seemed to show a slight inhibition of remote forward associations. When the poetry was used as learning material, the results showed a very definite inhibition of remote forward associations.

We may recall that, according to the principle of remote forward association, the strength of the forward associations decreases with the number of intervening items. Our tests were made with lists of words derived from every 2d word in the logical material, and this situation should afford the best possible opportunity for the formation of remote forward associations. Since the general results of the experiment are *negative*, the conclusion follows that, *when a series such as 1-2-3-4-5-6-7-8-9-10 is so learned that the items can be reproduced readily and in the proper order, this learning does not result in the formation of remote forward associations, and it does not facilitate the later learning of the items in the order 1-3-5-7-9 or 2-4-6-8-10. The results of the experiment indicate that the process of learning a series of items is not general, remote, or indirect, but very specific in its nature.*

Specific Serial Learning

From the results of the present experiment on remote forward association and from the results of the former experiment on backward association, we have reached the conclusion that *when a series of items such as 1-2-3-4-5-6-7-8-9-10 is well learned and practised until the recitation is very smooth, the forward associations between the adjacent items are the only associations that are formed.* We have shown that the associative process in serial learning is very specific in its nature, rather than haphazard, remote, or indirect. Although this conclusion may at first seem very simple, the implications are of special and wide significance for the psychology of learning. The remaining portion of this paper will show the relation between this fact and the broader problems of association and learning.[5]

[5] See E. Claparède, *L'association des idées*, 1903, especially pp. 120–130; H. Ebbinghaus, *Grundzüge der Psychologie* (2d ed.), 1905, 644–676.

In an article on the history of memory, Burnham writes that Plato, after referring to the block of wax in the mind of man, "calls it a 'waxen figment' and substitutes for it the figure of the aviary of all kinds of birds—'some flocking together apart from the rest, others in small groups, others solitary, flying anywhere and everywhere.' This receptacle is empty when we are young. The birds are kinds of knowledge. Learning is the process of capturing the birds and of detaining them in this enclosure." [6] Something similar to Plato's figure of the aviary of birds may still be found in several modern writers, especially among the philosophers. Much has been written about the *random character* of psychological processes, as well as the *casual behavior* of animal spirits. Our results show that in the limited problem of serial association *the learning process may be highly specific in nature.*

In summarizing Herbart's views on serial association, Burnham declares that "A series of presentations is reproduecd in consciousness in the same order in which it was originally perceived, for this reason: *each successive member of a series blends with all the other members of the series that are above the threshold of consciousness*,[7] with an energy proportional to the vividness of the latter. For example, in series *a*, *b*, *c*, *d*, suppose that the presentive activity of *b* is at a maximum and *a* is sinking, when *c* appears: then *c* will combine closely with *b*, lightly with *a*. When *d* appears, the presentive activity of *c* is at a maximum, *b* is sinking, and *a* is fading from consciousness; consequently *d* combines closely with *c*, lightly with *b*, very lightly with *a*; conversely *a*, which combined closely with *b*, combined lightly with *c*, and very lightly with *d*. Hence, after the series has vanished for a time from consciousness, and the initial member is in some way reproduced, *a* recalls *b* with an energy greater than that with which it recalls *c*, and so on; thus the series is reproduced in its original order." [8] Our results agree with Herbart's proposition concerning the

[6] W. H. Burnham, Memory, historically and experimentally considered, *Amer. J. of Psychol.*, 1888–1889, 2, 40–43.

[7] Italics not in the original.

[8] Burnham, *op. cit.*, p. 248.

order of reproduction, but we are doubtful about the process of "blending" and "combining." [9] When the series *a–b–c–d* is well learned and then repeated a large number of times, *the various items do not blend or fuse with each other.* We have shown that there is an association from *a* to *b*, but none from *b* to *a*,—and there is not even a slight tendency from *a* to *c* or from *a* to *d*. It is difficult to conceive of any two physical processes combining or blending with each other.[10]

Höffding formulated the law of association as follows: "For however many different sensations and ideas may come simultaneously, or in immediate succession, into our consciousness, they neither are nor remain quite separate. They are all embraced by one and the same consciousness, through whose activity they have arisen. The manner in which they act upon one another and are combined, is determined by the form and direction taken by the synthetic activity of consciousness at the given moment. On the other hand, they react, each one of them, upon the general condition of consciousness. Now when one of these sensations or ideas is renewed and brings the others with it, what really operates is the tendency to reawaken the general state, or the general activity, to which all these ideas belonged. The innermost basis of all association of ideas should thus be looked for in the unity which is present in every mental state and every mental activity, and which stamps all simple elements with a common characteristic. From this point of view the association between the parts and the whole would be the typical form of all association. This fundamental law of all association of ideas might be called the *law of totality*." [11] According to Hamilton's law of redintegration, "thoughts or mental activities having once formed parts of the same total thought or mental activity, tend ever after immediately to suggest each other." [12]

The results of our experiments are opposed to Höffding's law of totality and Hamilton's law of redintegration in so far as the statements of these authors have reference to well-

[9] See M. Bentley, A critique of 'fusion,' *Amer. J. of Psychol.*, 1903, 14, 60–72.

[10] We are not concerned in the present study with perception.

[11] H. Höffding, *Outlines of psychology*, 1891, 126, 159–160.

[12] See B. Edgell, *Theories of memory*, 1924, 83–90.

ordered serial association. After the series *1-2-3-4-5* has been learned and practised until the recitation is very smooth, item *3*, for example, does *not* tend to call up the whole series of responses, as Höffding and Hamilton would say. Item 3 tends to call up item *4*. Although *3* and *4* may be looked upon as contiguous stimuli, they do *not* tend to evoke a single joint response. Item *3* tends to evoke item *4*, and *4* tends to evoke *5*. Our study indicates that learning the series *1-2-3-4-5* does not in a strict sense result in the formation of a "higher unit." Item *1* evokes *2*, *2* evokes *3*, *3* evokes *4*, *etc.*, but *5* certainly does not tend to evoke *1*, *2*, *3*, *4*, and *5*. As far as well-ordered serial learning is concerned, we can subscribe to the claim made by Thorndike and Woodworth, in discussing the doctrine of transfer of training, that "The mind works in great detail." [13]

Law of Contiguity.—According to William James, "When two elementary brain-processes have been active together or in immediate succession, one of them, on recurring, tends to propagate its excitement into the other." [14] In the series, *1-2-3-4-5-6-7-8-9-10*, any two adjacent items, as, for example, *5* and *6*, have been "active together or in immediate succession" a very large number of times. There is a strong association from *5* to *6*, but *none* from *6* to *5*. The law of contiguity, as it is stated by many writers, tells *nothing* about such an important matter as the direction of the association between the two responses. It is evident therefore that some other factor, such as the mechanism of learning, is highly significant besides the mere *contiguity* of the two responses. Some authors, especially physiologists, still speak of contiguity as *the* law of association; but some psychologists have recognized that it is perhaps only correct when stated negatively. *If the law of contiguity is only true as a negative statement, then it is not a useful formulation of the law of the formation of associations.* Everybody knows that if *A* and *B* are now connected there must have been some relation or other between them at

[13] E. L. Thorndike and R. S. Woodworth, The influence of improvement in one mental function upon the efficiency of other functions. *Psychol. Rev.*, 1901, 8, 247–261, 384–395, 553–564. See also Thorndike's *The psychology of learning*, 1913, 350–357.

[14] W. James, *The principles of psychology*, vol. I, 1890, 566.

some time in the past. *The law of contiguity ignores the very factors which are important in association, and obscures the important problem of the mechanism of learning.* Some would say that associations are formed between successive rather than between simultaneous activities. It then becomes a question of *how much* time may elapse between the two responses; and the word "contiguity" begins to lose its meaning and whatever value it may have possessed.

Our results on backward association have the advantage in this connection of being very clear-cut and definite, and the implications for the older law of contiguity are unmistakable. If our negative conclusion in regard to backward association is justified, the law of contiguity must be rejected.[15]

Drainage Theory.—Our results are in agreement with James' drainage theory [16] in so far as he meant to say that when a series of items is learned the associations are in the forward direction and between adjacent items. James said much more than this, however; and in another place he accepted Ebbinghaus's results on remote association.[17]

Law of Resolution.—This law states that if a series of physiological states *A–B–C–D* is repeated several times, there is a tendency for condition *A* to pass at once to condition *D*.[18] Jennings states that "It seems probable that the same series occurs as before, save that conditions *B* and *C* are now passed rapidly and in a modified way, so that they do not result in a reaction, but are resolved directly into *D*." [19] Most of the experimental data on serial learning shows of course that *A* is not followed by *D*; but the question may still remain as to whether there is not some *tendency* for this to occur. If the law of resolution obtains, it would be expected that the continued repetition of *A–B–C–D* would alter the relation between

[15] Cf. W. McDougall, *Physiological psychology*, 1905, 118–139.

[16] James, *op. cit.*, vol. II, 1890, 579–592.

[17] *Op. cit.*, vol. I, 677–679. See also H. Cason, The physical basis of the conditioned response, *Amer. J. of Psychol.*, 1925, 36, 371–393; G. T. Ladd and R. S. Woodworth, *Elements of physiological psychology*, 1911, 577–580, 611–625.

[18] See H. S. Jennings, *Behavior of the lower organisms*, 1906, 289–292, 314–327, 333–335; E. L. Thorndike, *The original nature of man*, 1913, 185–192.

[19] *Op. cit.*, p. 290.

A and *D*, so that an association could, on a later occasion, be formed more easily from *A* to *D*. Our results show very clearly that these remote forward connections are not formed or strengthened. The continued repetition of *A–B–C–D* does not "decrease the resistance" from *A* to *D*, and it does not facilitate the later formation of an association from *A* to *D*.

The miscellaneous references given above have been selected partly at random, in order to show the relation between our results and several well-known concepts of association and learning. It should be apparent from our discussion that the *mechanism of learning* is a very important consideration.[20] It may be admitted that the theory of remote association seems quite plausible from an *a priori* standpoint; and it is difficult to conceive of a physical mechanism which could take care of the very specific serial associations which we have described, especially when one considers the complexity of some learning situations.[21]

Ebbinghaus claimed that more than twice as much time is required to learn a series of 200 items than one of 100 items;[22] and explanations have been offered partly in terms of the theory of remote association. It seems from our results, however, that if all of the conditions of learning could be kept quite constant and uniform—which of course is never possible —*a series of 200 items would be learned in just about twice the time required for 100 items.* The relative amount of material retained after a given interval of time should also be the same in the two cases.[23]

In all of the procedures which we employed an effort was made to limit the learning process to the formation of associa-

[20] In the learning situations which were involved in our experiments, "pleasure" really did not seem to "stamp in" the "successful" response! We have stated our views on the *conditioned response* elsewhere; and references to this theory of learning may be found in a general review of the literature by the writer in *Psychol. Bull.*, 1925, 22, 445–472.

[21] See K. Dunlap, The biological basis of the association of ideas and the development of perception, *Psychobiol.*, 1920, 2, 29–53; J. B. Watson, *Behavior, an introduction to comparative psychology*, 1914, 251–276.

[22] H. Ebbinghaus, *op. cit.*, pp. 46–61.

[23] *Cf.* E. A. M. Gamble, A study in memorising various materials by the reconstruction method, *Psychol. Monog.*, 1909, 10 (no. 43), 59–89, 150–210.

tions in the forward direction, from *A* to *B*, from *B* to *C*, *etc.* The learning material was composed of a large number of items, and it does not seem plausible to believe that many associations were formed with the *absolute position* of the items. Each response or series of responses seemed to sprout out of those which immediately preceded.

Before concluding this discussion, let us suppose that a series *1-2-3-4-5-6-7-8* represents an incident which occurred in the past. Now let items *3* and *4* or the connection between them suffer "atrophy through disuse" or decline, as the result of some other cause. It may be then that the items *5-6-7-8* cannot be recalled. An appropriate stimulus can evoke *1* and *2*; but since *3* and *4* are in a poor condition, the series *5-6-7-8* is gone. In order to revive *5-6-7-8*, we can think "about" and "around" the items *3* and *4*, and in this way encourage their revival,—so that the series *5-6-7-8* may unroll itself quite unexpectedly. The lost items may of course remain in oblivion, or, as sometimes happens, they may have to be revived periodically. Many of our responses are not linked together in a single series, but are arranged in a network extending in many directions. For this reason, various stimuli may be adequate to revive a series that has been forgotten. Although there are numerous complications, the adequate stimuli are frequently very *specific*.[24]

In the discussion which has preceded, we have frequently referred to the "items" of a series and the associations which are formed between the "individual responses." Of course these elements do not operate in a manner which would indicate that they are completely separated from other processes. Although we are seldom stimulated through just one receptor at a time, it frequently happens that one stimulus or one response occupies the center of the stage at a given moment. The interesting "configuration" also does not exist completely separated from other functions.[25] The "configuration" is

[24] See E. B. Titchener, *A text-book of psychology*, 1910, 384–389.

[25] The writer spent a good many hours observing the 95 subjects while they were learning the rote material. Practically all of the subjects made regular movements of the mouth, tongue, throat, and frequently even the hands, while they were learning the material. It is difficult to believe that these movements did not form a part of the learning process.

similar to our response in being merely *a part* of a still larger and more complicated situation. Many of the criticisms of "associationism" cannot be justly applied to our experiments, because we employed several types of material, because the procedure was varied frequently, and because serial learning—or habit formation—is a very common occurrence in everyday life.[26]

References

Bentley, M. A critique of 'fusion.' *Amer. J. Psychol.*, 1903, **14**, 60–72.

Burnham, W. H. Memory, historically and experimentally considered. *Amer. J. Psychol.*, 1888–1889, **2**, 39–90, 225–270, 568–622.

Carr, H. A., & Freeman, A. S. Time relationships in the formation of associations. *Psychol. Rev.*, 1919, **26**, 465–473.

Cason, H. The concept of backward association. *Amer. J. Psychol.*, 1924, **35**, 217–221.

Cason, H. Specific serial learning; a study of backward association. This Journal, 1926, **9**, 195–227.

Cason, H. The physical basis of the conditioned response. *Amer. J. Psychol.*, 1925, **36**, 371–393.

Cason, H. The conditioned reflex or conditioned response as a common activity of living organisms. *Psychol. Bull.*, 1925, **22**, 445–472.

Claparède, E. *L'association des idées*, 1903.

Dunlap, K. The biological basis of the association of ideas and the development of perception. *Psychobiol.*, 1920, **2**, 29–53.

Ebbinghaus, H. *Memory, a contribution to experimental psychology*, 1885. (Tr. by Ruger and Bussenius.)

Ebbinghaus, H. *Grundzüge der Psychologie* (2d ed.), 1905.

Edgell, B. *Theories of Memory*. Univ. of Oxford Press, 1924.

Frings, G. Ueber den Einfluss der Komplexbildung auf die effektuelle und generative Hemmung. *Arch. f. d. ges. Psychol.*, 1914, **30**, 415–479.

Gamble, E. A. M. A study in memorising various materials by the reconstruction method. *Psychol. Monog.*, 1909, **10** (no. 43).

Höffding, H. *Outlines of psychology*, 1891.

James, W. *Principles of psychology*, 1890.

Jennings, H. S. *Behavior of the lower organisms*, 1906.

Ladd, G. T., & Woodworth, R. S. *Elements of physiological psychology*, 1911.

McDougall, W. *Physiological psychology*, 1905.

Michotte, A., & Fransen, F. Note sur l'analyse des facteurs de la mémorisation et sur l'inhibition associative. *Ann. de l'institut sup. de philos.* (Université de Louvain), 1914, **3**, 503–549.

Müller, G. E., & Pilzecker, A. Experimentelle Beiträge zur Lehre vom Gedächtniss. *Zsch. f. Psychol. u. Physiol. d. Sinnesorgane*, Ergänzungsband 1 (1900), 1–288.

[26] For a practical discussion of habit, see S. H. Rowe, *Habit-Formation and the Science of Teaching*, 1909, 44–68.

MÜLLER, G. E., & SCHUMANN, F. Experimentelle Beiträge zur Untersuchung des Gedächtnisses. *Zsch. f. Psychol. u. Physiol. d. Sinnesorgane*, 1894, **6**, 81–190, 257–339.

NAGEL, F. Experimentelle Untersuchungen über Grundfragen der Assoziationslehre. *Arch. f. d. ges. Psychol.*, 1912, **23**, 156–253.

ROWE, S. H. *Habit-formation and the science of teaching*, 1909.

THORNDIKE, E. L. *The original nature of man*, 1913.

THORNDIKE, E. L. *The psychology of learning*, 1913.

THORNDIKE, E. L., & WOODWORTH, R. S. The influence of improvement in one mental function upon the efficiency of other functions. *Psychol. Rev.*, 1901, **8**, 247–261, 384–395, 553–564.

TITCHENER, E. B. *A text-book of psychology*, 1910.

WATSON, J. B. *Behavior, an introduction to comparative psychology*, 1914.

WOHLGEMUTH, A. Ueber die Richtung der Assoziationen. *Ber. über d. V. Kongress f. exper. Psychol.*, 1912, 284–287.

WOHLGEMUTH, A. On memory and the direction of associations. *Brit. J. Psychol.*, 1913, **5**, 447–465.

AN OBSERVATIONAL STUDY OF ANGER

BY GEORGINA STICKLAND GATES

Barnard College, Columbia University

In attempting to describe the emotional reactions of adults the psychologist is occasionally embarrassed by the lack of concrete data upon the frequency, causes, usual conditions, reactions and impulses of the emotional crisis. How often, for example, does the average adult experience anger? What is the most frequent impulse during this state? How may the causes of anger be classified and described? At what time of day is anger most likely to occur? Do the after-effects always give rise to an unfavorable bodily condition? All these are questions for which we have at present no adequate answer, but to which answers may, in spite of their inevitably humorous aspect, add something by way of our knowledge of human nature.

A first attempt at answering these questions in quantitative fashion is made by the present study. The method of introspection and report immediately following an experience is used here in contrast to the method of the questionary previously employed by Hall.[1] Our data are treated statistically rather than descriptively, as in Hall's investigation and in observations made by twelve subjects of Richardson,[2] and they are considered, in the main, from the point of view of the circumstances and reactions involved, rather than from that of the consciousness present. In our study the data were recorded as soon as possible after the actual occurrence. All the individuals reporting were of approximately equal age and like social status, and subject to similar environmental conditions.

Fifty-one women students, members of two introductory

[1] G. S. Hall, A study of anger, *Amer. J. of Psychol.*, 1899, 10, 516–591.

[2] K. F. Richardson, The psychology and pedagogy of anger, *Educ. Psychol. Monog.*, no. 19.

courses in psychology, agreed to record all their experiences of anger or extreme irritation during one week. Each subject graded each anger experience on a scale from 5 to 1, calling 5 the most extreme anger ever felt by the subject, 1 the least degree of anger ever experienced, 3 intermediate between these extremes, and 4 and 2 in their respective places. She recorded the date, exact time of day (noting also whether the emotion occurred shortly before or after a meal); she graded her condition before the situation causing the anger arose on a like scale, 3 representing her usual state of health, 5 an exceptionally good, 1 an extremely poor condition, *etc.*, and added a descriptive word (as tired, cold or excited); she described the exact cause as well as she could determine it, the responses made, the impulses experienced but not yielded to, and any after-effects noted.

Before the experiment was begun each subject estimated her own "degree of irritability," *i.e.*, she guessed whether or not she was easily angered. This scoring was done also in a scale of 5 to 1, 5 representing the most easily angered person of her acquaintance, 1 the least easily aroused, with herself ranked accordingly.

The reports of the 51 Os yielded 145 cases of anger; 6 of degree 5, 23 of degree 4, 44 of 3, 49 of 2, and 23 of 1. Bearing in mind the caution that this analysis is applicable only to these particular subjects under these conditions, and that it is further limited by the difficulties of introspection under which inexperienced observers labor, certain conclusions to be described briefly may be drawn.

I. *Frequency.*—The instances of anger reported for the week are distributed among the Os as follows:

Instances reported	0	1	2	3	4	5	6	7	8	9	10
Os reporting	6	12	4	9	14	2	2	0	0	0	2

One's own angers may profitably be compared with this presumably normal incidence in young women living in a college community.

II. *Accuracy of Self-Rating.*—In Table I the individuals making the observations have been divided into four groups,

TABLE I

COMPARISON OF FREQUENCY AND AVERAGE DEGREE OF ANGER REPORTED WITH SELF-RATING BEFORE OBSERVATIONS WERE MADE

Self-Rating	Average Number Occurrences Reported	Average Degree of Anger	Number of Cases*
4	3.81	2.69	11
3	2.71	2.53	17
2	2.66	2.19	15
1	2.00	2.10	3

* Five individuals were not present when the rating was made.

representing their original self-ratings (as described above), and the average frequency and degree of anger have been computed in each group. There appears to be a very good correspondence (though the differences between the groups are small) between the original judgment and both the frequency and the degree of the emotion experienced. Those who rated themselves as 4 reported 3.8 occurrences; those rated 3, reported 2.7; and so on. The average degree of anger varies similarly with the original judgment. Since it is extremely improbable that these ratings influenced the number and degree of emotions reported (for in most cases the subjects were ignorant of the records of others and so could not know whether they were describing a relatively large or small number of cases of anger) it seems safe to conclude that, in general, the judgments were based on some real though possibly ill-defined knowledge of conditions or characteristics.

III. *Time of Occurrence.*—Tables II and III show the frequency of occurrence of different degrees of anger on the various days of the week and at the different times of the day. Friday, Saturday and Sunday average 26⅓, which contrasts with the 15¼ average for Monday, Tuesday, Wednesday and Thursday. Cases of anger seem to be of much more frequent occurrence on those days which are holidays or (as was the case with Friday for most of our students) half-holidays. Even when we omit Friday a similar difference appears. Various reasons were suggested by the subjects for this wide difference, varying from opinions that the subjects were more

tired at the end of the week to suggestions that one is more likely to be angered when engaged in social as compared with intellectual pursuits. The list of causes will show, as I think, that opportunity rather than predisposition is the main reason for this difference. There is no evidence of the existence, in this group, of a "blue" or irritable Monday.

TABLE II

NUMBER OF CASES OF ANGER OCCURRING ON EACH DAY

Degree	Friday	Saturday	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday*
V....	1	0	2	0	0	1	2	0
IV....	5	3	4	2	2	4	2	1
III....	11	10	5	3	4	6	3	2
II....	8	9	10	6	4	5	6	1
I....	5	6	0	4	5	2	0	1
Total	30	28	21	15	15	18	13	5

* Cases of anger on this day were reported only when occurring before 10 A.M.

TABLE III

NUMBER OF CASES OF ANGER OCCURRING AT VARIOUS TIMES OF THE DAY

Degree	Less than One Hour Before Breakfast	At Breakfast	Less than One Hour After Breakfast	Middle of Forenoon (About two hours)	Less than One Hour Before Lunch	At Lunch	Less than One Hour After Lunch	Middle of Afternoon (About three hours)	Less than One Hour Before Dinner	At Dinner	Less than One Hour After Dinner	In the Evening (About three hours)	After Retiring
V.....	0	0	0	0	1	0	0	1	3	0	0	0	1
IV.....	2	0	0	1	4	0	1	5	5	0	1	4	0
III.....	3	0	5	3	5	0	3	6	8	3	3	3	2
II.....	6	1	7	5	6	0	2	2	7	0	2	10	1
I.....	4	0	5	0	2	0	1	2	0	2	2	5	0
Total	15	1	17	9	18	0	7	16	23	5	8	22	4

Our data concerning the time of day at which anger occurs most frequently confirm popular opinion. The high point is reached just before dinner; the hour before meals standing to the hour immediately following as 18.7 to 10.7. If we omit the breakfast period from our calculations (since there are an

unusually large number of cases occurring immediately after this meal, due to delays encountered in "getting to college"), the averages for before and after meals are 20.5 and 6.0. The more violent cases of anger are also more likely to occur just before lunch or dinner; 17 per cent. of the cases of degrees IV and V occur before lunch; 28 per cent. before dinner; whereas only about 12 per cent. of I and II are reported before lunch and 10 per cent. before dinner; $\frac{2}{3}$ of degree V occur before a meal. The afternoon seems to be the favored time for experiences of this sort as compared with the periods between breakfast and lunch and during the evening.

IV. *Condition before Anger Occurred.*—A slightly greater number of individuals reported themselves as feeling worse than usual before the occasion of the anger than rated themselves as feeling better than usual; although we find a fairly even distribution (Table IV) of conditions. There appears to be no difference in average condition reported between the two greatest and the two least degrees of anger. When we consider the descriptive terms which were given in about two thirds of the cases (Table V) we find 65 reports of poor conditions (tired, sleepy, tired and hungry, hungry, headache, or cold, irritable, frightened) and 22 reports of good conditions (excited, "feeling good," relaxed, asleep). The relative excellence of the condition does not, however, vary with the degree of anger. Evidently it is not the condition of the organism, as judged in this way at least, which is the main determinant of the degree of anger.

TABLE IV

CONDITION BEFORE SITUATION CAUSING ANGER OCCURRED

Degree	Condition Rated					
	1	2	3	4	5	Average
V	0	3	2	0	1	2.8
IV	1	8	3	7	4	3.2
III	0	20	10	10	4	3.0
II	2	14	19	10	4	3.0
I	0	6	9	6	2	3.2
Total	3	51	43	33	15	

TABLE V

DESCRIPTION OF Os' CONDITION BEFORE THE ANGER SITUATION AROSE

State	Degree					
	V	IV	III	II	I	Total
Tired	1	5	13	7	7	33
Sleepy	0	0	2	4	2	8
Tired and hungry	1	1	3	2	1	8
Hungry	1	1	0	2	0	4
Headache or cold	1	1	2	2	1	7
Irritable, worried		1	2		1	4
Frightened				1		1
Excited, happy, gay	0	1	6	1	1	9
Good	0	2	1	1	2	6
Relaxed, contented	0	0	1	3	0	4
Asleep	0	1	1	1	0	3
Not given	2	10	13	25	8	58

V. *Duration.*—The experiences of anger varied in duration from less than five minutes (in 50 cases) to one day (in 2 cases) with the median between 10 and 20 minutes (Table VI). The more severe the anger, the longer, in general, it lasts.

TABLE VI

THE DURATION OF ANGER

Degree	5 min or less	10 to 20 min	½ to ¾ hr	1 to 1½ hr	2, 3, 4 hr	½ Day	1 Day	Median (min)
V	0	1	2	1	1	1	0	45 to 60
IV	4	8	1	5	1	2	2	10 to 20
III	14	19	4	5	2	0	0	10 to 20
II	14	21	6	5	3	0	0	10 to 20
I	18	2	2	1	0	0	0	5 or less
Total	50	51	15	17	7	3	2	

VI. *Causes.*—The concrete causes of anger given by the subjects included unjust accusations, insulting or sarcastic remarks, contradictions, criticisms or scoldings, unwelcome advice, others "knew too much," "being bossed" by parents or friends, being teased, work left for subject to do, being kept waiting by friends, "not invited to the party," "being shoved," stepped on, hat pushed off, seat taken, the sight of others being rude or unjust, disobedience of children, mother, sister or friend refused request, spilling the ink, being locked

out, wrong number, locker, radio or typewriter wouldn't work, umbrella, fountain pen or money lost, clothes injured, glasses or watch broken, hair wouldn't stay up, lights went out, fumbled in dressing or sewing, dog refused to obey, elevator or bus slow, study or sleep interrupted, store not open, physical pain and thwarted hunger.

A certain amount of difficulty was experienced in classifying these causes. Thwarting of some activity, either instinctive or learned, which is under way, is frequently described in discussions of anger as the main cause for the emotion. Thwarting of self-assertion appeared so often in various guises in our own records that we gave this impulse a separate classification. In this grouping we found various methods by which the "mastery impulse" might be thwarted, and these corresponded with a description of three of the forms of self-assertion given by Woodworth.[3] We then placed under thwarting of self-assertion (1) anger caused by the necessity for a "defensive reaction to persons" or resisting domination as in the case of an insulting remark or contradiction; (2) that stimulated by a resisted "aggressive reaction to persons" or thwarted efforts to dominate others as in the case of refusal of requests by others; and (3) that caused by the arousal of a "defensive reaction to things," or efforts to overcome obstruction, as when the radio refuses to work.

Apart from these emotions which seem to be caused primarily by the thwarting of this mastery impulse, are those in which the element of interrupted habitual or instinctive activity of some more specific kind is more evident. The anger caused by interrupted study or sleep or by inability to enter a desired athletic competition is possibly different in its causation from that stimulated by insults, contradictions, scoldings, disobedient children, or even spilling the ink.

In many cases the causes overlap; in these the experimenter has chosen the most probable group for classification. "Delays" occur in three places, those where the anger seems to be aroused by injury to pride (as being kept waiting by one's friends), which is placed under defensive reaction to

[3] R. S. Woodworth, *Psychology; a study of mental life*, 1921, 161–166.

persons; those where anger is directed at the offending elevator or bus, under defensive reaction to things; and those where the delay is annoying because a desired activity is postponed under definite activities thwarted. Table VII shows (in percentages) the cases of anger attributable to each cause. In 115 of the 145 cases, persons, rather than things, are the main cause of anger. Thwarting of the impulse of self-assertion accounts for 92 of the experiences; and it is probable that other cases have, as one of the stimulating elements, the obstruction of this same impulse.

TABLE VII

DEGREE OF ANGER AS CAUSED BY FOUR KINDS OF OCCASION (IN PERCENTAGES)

Degree	Thwarting of Instinct of Self-Assertion			Thwarting of Other Activities
	Defensive Reaction to Persons	Aggressive Reaction to Persons	Defensive Reaction to Things	
V....	67	0	0	33
IV....	39	4	22	35
III....	45	5	11	39
II....	25	14	16	45
I....	31	0	52	17

The various degrees of anger differ perhaps more markedly in their causes than they do in any of the other circumstances studied. About half (52%) of the cases of degree I are referable to obstruction by things; while none of the cases of degree V is caused by this kind of thwarting. A third (31%) of the cases of the mildest degree of anger are caused by the domineering attitude of other persons, while two-thirds of degree V are so caused. Probably we have a genuine difference here. Not only do persons cause anger much more frequently than do things, but the emotion experienced is much more likely to be violent in the former than in the latter case.

VII. *Reactions During Anger.*—An amusing picture of the responses of the young college girl to offending persons or objects and also an enlightening commentary on the inefficiency of human instinctive reactions are suggested by Table VIII,

TABLE VIII

RESPONSES MADE DURING ANGER

	Degrees					
	V	IV	III	II	I	Total
Gross Bodily Responses Directed at Offending Object:						
Excited talking or angry exclamation	2	2	10	13	5	32
Angry, sarcastic, sulky retort	0	6	8	9	3	26
Restless behavior (*e.g.*, pacing, tossing in bed, shifting in chair)	2	2	7	7	2	20
Refusal to speak to or look at the offender	0	7	3	5	3	18
Violence to inanimate objects	1	2	2	4	1	10
Sudden exit from room	1	0	1	4	1	7
Grimace or glaring and staring at offender	0	0	0	5	0	5
Violence to offender (slap, shake)	0	1	0	2	0	3
Refusal of food	1	0	0	1	0	2
Pleasant reply	0	0	1	1	1	3
Sum	7	20	32	51	16	126
Expressive Movements:						
Unpleasant facial expression (frown)	2	8	6	10	4	30
Biting of fingers or lips	0	3	6	2	2	13
Clenching teeth or hands	4	2	2	1	3	12
Body tense	3	4	0	2	1	10
Stamping foot	1	1	1	3	2	8
Tears in eyes	4	0	1	1	2	8
Eyes flashed, stared, popped	0	1	0	1	1	3
Sum	14	19	16	20	15	84
Activities of Sympathetic System and Adrenal Glands (Mainly):						
Gasp, heavy or rapid breathing	1	2	7	11	7	28
'Hot' feeling	3	5	5	4	2	19
Flushing	0	1	6	5	1	13
Fast heart beat	2	0	8	2	0	12
Swallowing, choked feeling	2	0	2	0	0	4
Nausea or 'sinking' in stomach	1	2	1	0	0	4
Trembling, weak feeling	1	0	0	1	2	4
Cold hands, dry lips	0	2	0	0	0	2
Loss of desire to eat	0	0	2	0	0	2
Headache, dizziness	0	0	2	0	0	2
Sum	10	12	33	23	12	90

which enumerates the reactions made. The soft answer occurs, it is to be noted, in less than one per cent. of the cases (and these mild), a result contrasting with that of Richardson who found 18 per cent. of his reactions to be of the "contrary" type. Shaking or slapping the offender is reported in only 3 of our 145 cases. The most frequent bodily responses described are—as any careful observer of human nature might guess—verbal, "talking excitedly to others or making an angry exclamation;" "angry, sarcastic, sulky retort;" and restless behavior of various kinds, pacing, tossing in bed, shifting

about in the chair, continuous singing, and, that refuge of the offended maiden, refusing to speak or look at the offender. An unpleasant facial expression is described 30 times, and some change in breathing almost as frequently.

A larger number of responses are reported in proportion to the number of cases in the more violent experiences of anger, and this is especially true of some of the responses which are caused mainly by the activities of the sympathetic system, as nausea, hot feeling, and cold hands.

VIII. *Impulses Experienced.*—We might expect to find the impulse most frequently reported by the subjects to be the primitive desire to perform those acts for which anger is supposed to be a preparation, *i.e.*, to do physical injury to the offender; and this tendency is often described (see Table IX). More frequent, however, is the learned impulse for which none of the internal responses appears to be essential, to make a verbal retort, as to "tell her what I thought about her," to

TABLE IX

IMPULSES FELT DURING EXPERIENCE OF ANGER

	Degrees					
	V	IV	III	II	I	Total
To make verbal retort ("tell her what I thought," scold, say "I hate you")	3	8	16	20	6	53
To do physical injury to offender (slap, pinch, shake, strike, choke, push, step on, scratch, shoot, beat up, throw out of window, kill, tear to pieces, throw something at, spank	2	7	14	15	2	40
To injure inanimate objects (tear up, bang down, slam door, kick, knock on floor)	2	2	3	7	6	20
To run away, leave room	2	2	4	3	1	12
To cry, scream, swear	1	1	4	2	2	10
To "do what I wanted to"	0	0	4	3	2	9
To give up efforts	0	3	3	1	1	8
To tell others about it, or appeal to authority	0	1	2	1	0	4
To retaliate in kind	0	1	0	2	0	3
To blame some one else	0	0	1	0	2	3
To be perverse	0	0	0	0	1	1
To laugh	0	1	0	0	0	1
To refuse to eat	0	0	0	1	0	1
To tear my hair	0	0	0	0	1	1
Sum	10	26	51	55	24	166

insult and to scold. One wonders if a similar result would be obtained from a census of the impulses of the male. A desire

experienced 12 times is "to run away, to leave the room;" which seems, according to the comments of the subjects, to be either an expression of displeasure or a wish to escape from the situation.

Comparing the various degrees of anger, we find that more impulses are reported in connection with degree V than with degree I, and that the more violent tendencies are more frequently described as occurring with the highest degrees of emotion. In these extreme cases, however, relatively mild impulses are also experienced, and, contrariwise, the more injurious are at times reported in connection with the lesser degrees of irritation.

IX. *Reactions Experienced after the Emotions have Passed.* —"No reaction" is reported in 28 per cent. of the cases (Table X), an 'unfavorable' reaction in 51 per cent., and a 'favorable' in 15 per cent. The most frequent after-effect is a feeling of irritability; next, shame or disgust with one's self; "tired" is reported 18 times and a feeling of satisfaction or triumph 11 times. More of the unfavorable reactions are

TABLE X

EXPERIENCE AFTER ANGER HAS PASSED

	Degrees					
	V	IV	III	II	I	Total
Irritable, out of sorts	0	6	6	8	6	26
Disgusted with self, ashamed, foolish	1	6	5	10	2	24
Tired	0	4	7	5	2	18
Increased dislike of person who caused anger	1	0	7	2	0	10
Sad, hurt, lonely	2	0	0	0	0	2
Felt "dumb"	0	0	1	1	0	2
Satisfaction, triumph, feeling of energy	1	1	2	3	4	11
Relief	0	3	1	0	1	5
Resignation	1	0	0	5	1	7
No reaction	0	3	15	15	7	40

reported as following the violent cases of anger; the favorable reaction is more likely to occur in connection with the milder degrees. "No reaction" is similarly reported more frequently in the description of the lesser experiences of irritation.

SUMMARY

I. The average number of experiences of anger during a week, is, for our 51 subjects, 2.85.

II. Each subject is able to estimate with a fair degree of reliability her position on a short scale running from most to least easily angered.

III. Anger occurs more frequently on holidays than on week days, and before, rather than after, meals.

IV. The bodily conditions preceding anger are more frequently poor than good; but there appears to be no correspondance between bodily condition and the intensity of the emotion.

V. The median experience of anger lasts from 10 to 20 minutes, the range being from less than five minutes to one day.

VI. Thwarting of self-assertion is more frequent in causing anger than is the obstructing of any other instinct or habit, and persons more frequently than things are provocative of the emotion.

VII. The reactions and impulses reported show a predominance of the verbal response.

VIII. The after-effect of the emotion is unfavorable in about half of the cases, the most frequent reaction being continued irritability, shame, or weariness.

IX. The more violent cases of anger are to be differentiated from the milder by their longer duration, greater number of reactions and impluses, unfavorable after-effects, and especially by a difference in cause, a defensive reaction to the domineering attitude of others being the chief factor in the more violent cases, an irritation caused by obstructing objects being often operative in the milder.

The investigator suggests the desirability of performing a similar experiment on male subjects so that facts might be obtained upon the possibility of sex-differences in temperamental traits. An inventory of causes, impulses, responses, and the conditions of anger would give a picture of the usual male reactions which might differ, to some extent, from the composite portrait presented here of the rage experienced by the young woman. It would be relatively easy, by means of similar studies, to demonstrate changes with age, experience or environmental condition.

VISUAL PERCEPTION IN THE WHITE RAT

BY GLENN D. HIGGINSON

University of Illinois

Ingenious researches into the performance of vertebrates, especially of the anthropoid apes, have recently served to revive a flagging interest in the experimental study of the animals. Matters of great import to comparative psychology, long accepted without question, have been challenged in a noteworthy manner. The work of Koehler [1] upon the higher apes is outstanding, not only in its scope but also in the character and the control of the experimental procedures followed. Koehler chiefly appealed, we may say, to the apprehensive functions of the animals under objective conditions which provided means of determining the ability of the ape so to extend a perceptual situation as to include a number of unrelated phenomenal items. The experimental setting approximated the customary environmental conditions, and the results consequently possess a peculiar scientific interest. In his interpretation of results, Koehler found much material exemplifying and substantiating, as he thought, the fundamental principles in the doctrine of configuration. It is not our purpose, however, either to discuss these investigations or their precise significance for the *Gestalttheorie*. We have referred to them because we find in the performances of the white rat something which we believe to be directly comparable to the activities of Koehler's apes.

According to the exponents of configuration, a perception is primarily characterized by its highly integrative and unitary aspects. A totality, not a sum of individual constituents, is the primary outcome of apprehension; the animal tending to perceive 'wholes' and not 'parts'; unified structures and not conglomerates. The 'parts' of these phenomenal unities

[1] W. Koehler, *The mentality of apes*, 1925.

gather their particular significance from the total *Struktur* of which they are members, membership depending upon the degree to which they partake of the character of the whole or contribute to a particular functional outcome. That is to say, there is in perception a wide extension of functional performance whereby various objects are phenomenally integrated through a common purpose or goal or under a common need. In this apprehensive ligation the value of one object rests primarily upon its induction into a total situation. Thus a short stick utilized by the ape in securing a longer one by which food may be obtained is incorporated into the perceptual totality only when it bears upon it the functional mark of the total *Struktur*, namely, the obtaining of food. The stick comes to be manipulated under a wide extension of a formerly restricted meaning; for the stick is no longer, in any sense, a mere stick or object: it is a precise means of advancing an objective or of attaining an end. Moreover, to indicate the functional significance of a given 'member,' operative among others in furthering the organism, we may cite the fact that almost any object may, in the situation suggested, come to have the same functional value as the original stick. There results, then, a partial or total submersion in the total and unified whole of the individuality which previously characterized the object. It is no more just 'that thing' but has become even as Kinnaman put it years ago "a thing-which-has-something-to-do-with-food-getting." The essential feature of this unifying process is that the apprehended relation of a particular member to the whole does not necessarily arise through a long sustained period of "trial and error." The induction may be quite sudden. In fact, we might say that the major contribution of these experiments consists in showing that, instead of an elaborate and sustained process leading to the final outcome, the 'solution' of a perceptual situation issues immediately. The goal, or end-result, even though fairly remote in time, functionates always as an integral part of a total *Gestalt*, determining the character of the animal's performance by bringing the members into a greater unity.

Without denying or affirming intelligence in the anthropoid

apes—which is, after all, largely a matter of definition—we will doubtless agree that, in common with many other infra-human and human organisms, the apes draw heavily upon visual perception as a major determinant of action. This fact must always be recognized when considering the matter of animal performance. Carr points out that the "process of perceiving is a constituent part of a larger act of adjustment and can be adequately comprehended only in terms of these wider relations." [2] Most cases of intelligent behavior can be explained, no doubt, in terms of an intimate integration of the executive functions with visual apprehension under widely extended instruction of the 'occasional' kind. The remarkable performances of Clever Hans and of the Elberfeld horses, for example, have been shown to be cases of the close fusion of action with visual perception. Slight visual cues, usually imperceptible to the casual spectator, served to release and to check the course of action. What is true for these animals holds for others as well, including as we believe, the white rat. Our purpose here is to show that the visual capabilities of the rat are greater than men generally recognize; and that the rat utilizes visual resources in orientation more than we are commonly led to believe.

In approaching this matter of visual perception in the rat let us ask whether the behavior of the rat indicates, in common with Koehler's apes, an extension of visual apprehension such as to initiate a sudden shift in the actional performance which is incapable of explanation by the usual conception of 'trial and error' and by the 'laws' of frequency and recency. Can the rat visually perceive to advantage a sudden change in its path which offers a major reduction in its usual course toward food? Can it, *e.g.*, in its quest for food see through a perceptual situation as the apes did? We believe that we have shown by experiments that rats do employ, in a novel manner, their apprehensive functions for adjustive purposes; that they are not, in any sense, automata, living machines controlled by kinæsthetic thrusts and precise neural patterns, as some have claimed, but instead organisms capable upon demand of sudden readjustments to significant environmental changes.

[2] H. A. Carr, *Psychology; a study of mental activity*, 1925, 133.

To test the visual capabilities of the rat we made the following modification in a standard circular maze. At the entrance into alley 3 we placed a snugly-fitting door. During each of 100 runs the rat coming down alley 2 found a smooth wall to its right. If it ran to the end of the *cul-de-sac*, it found on its left, upon returning, the door open into alley 3; if, however, it failed to run the full distance to the end of the blind alley the door remained closed. The rat under no circumstance found the door open on its right. At the end of 100 trials, during which the animal each time went through the door into alley 3, turning always to the *left* (after running six feet out of its way), the animal was again introduced into the starting box and the alley 3 door, previously closed, was opened. Now one hundred runs necessarily meant a high degree of training for every animal, since the average rat learns the maze in something less than twenty-five trials. We expected that, with this amount of repetition, the animal would run, in spite of the open door, to the end of the alley and back again and so into alley 3 to the left. Curiously enough but four of the nine rats did and these four immediately shifted. The remainder exhibited the type of performance which most interested us. They stopped suddenly and without interference ran the remainder of the maze correctly, thus dropping at once six feet from the previous pathway and turning in a different manner. Apparently the sudden elimination of this extra six feet no more affected the total run than the addition, through error, of an occasional blind alley affects the remainder of the run of any rat acquainted with the maze. We know that even the best trained animal will upon occasion run into a blind alley or suddenly turn back upon its course only to finish perfectly the remainder of its run. Now this sudden elimination of six feet from the total pathway set up under continued repetition is wholly inexplicable in the usual categories of 'kinæsthetic patterns' and 'frequency and recency' of performance. Since the animals had never entered the door turning to the right, they could not have done so when the door was open, if we take frequency as the basis of performance; and if kinæsthesis ruled the animal the lifting

of the door should not have made any significant difference in performance; but since it did that fact must find explanation in visual terms. This employment of visual functions in the orientation of our subjects shows clearly that, for the particular demand made upon our animals, visual perception constitutes, under the conditions, an important functional resource of the organism.

It has long been widely asserted that in the acquisition of the maze visual factors are negligible in the rat; that kinæsthetic items constitute the vehicle by which the course of the animal is sustained; and that correctness of the course, including the elimination of errors, depends upon the frequency with which the true course is travelled in comparison with the number of times the animal enters a given blind alley. The rat thus becomes a mechanism wound up and needing only an introduction into the maze to release neural patterns which lead to a perfectly stereotyped performance. This way of regarding the rat may be traced particularly to the influence of Carr and Watson.[3] The studies from which spread the impression that rats are machine-like do not appear, as we find, to warrant such an assumption; to claim otherwise is to disregard individual differences at the risk of obscuring the facts. With the general interpretation which has developed from these studies the writer has little sympathy. He is unable to agree, for example, that the white rat is incapable of sudden readjustments to a total situation through the novel use of the perceptive functions. His opinion, moreover, finds clear evidence of a confirmatory character in the literature. Since the contributions to this matter are extensive, we shall consider only the more outstanding.

Small[4] was the first to claim that the rat does not utilize visual functions in the maze; he based his conclusion upon the nature of the rat's response to visual 'cues' introduced at critical points in the maze. But we must question whether

[3] H. A. Carr & J. B. Watson, Orientation in the white rat, *J. of Comp. Neur. & Psychol.*, 1908, *18*, 30; J. B. Watson, Kinæsthetic and organic sensations, *Psychol. Rev., Monog. Supp.*, 1907, *8*, 75.

[4] W. S. Small, Psychic development of the young white rat, *Amer. J. of Psychol.*, 1899, *2*, 63.

we may rightfully assume that, because the removal of the colored pegs used by Small did not interfere with the performance of the animal, visual perception was not employed. Must we infer that in the rat the total perceptual apprehension of a visual sort is hung upon a colored peg. This is not true of the human being under similar circumstances. The removal of a single object from a familiar situation does not result in a large shift either in perception or in action. Watson inferred from the effects of a progressive elimination of the functional receptors in the rat, as well as from a comparison of maze performances in the light and in the dark, that the animal in the maze drew little, if at all, upon vision. Carr, too, concluded from changes in the visual pattern of the maze and from introducing the rat into the maze at various places, as well as by lengthening and shortening the compartments, that the rats differed from the human subject in demanding kinæsthetic, instead of visual, cues. The rat in place of using vision, when a break occurs in the series of acts, gets a kinæsthetic cue "by traversing at random any unit of the maze." It is admitted, however, that normal rats make required spatial orientations much more quickly than blind animals. "Since two out of the eight animals (to quote Carr) made eight of the nine unquestionable immediate orientations we are willing to admit the possibility of the use of distance sense data in their case." [5] In this same study, we find that, out of 84 trials in which animals were released at random in a maze, the rat immediately oriented itself 17 times and ran correctly the remainder of the maze. Further to indicate that vision was here operative we find that the animals entered the *cul-de-sacs* only eight times in the 84 trials; although they were set down along the maze in such a way that kinæsthetic factors could not have been immediately utilized. From some unpublished studies of a similar nature conducted under the direction of Professor Bentley it was found that white rats immediately orientated themselves when introduced at various places into a maze with which they were familiar and completed their runs without interference. Since kinæsthesis had no

[5] Carr & Watson, *Orientation, etc.*, p. 37.

chance here to function, the inference was that visual resources were responsible for correctness of the animal's responses. Vincent[6] asserts that the eye of the rat is fairly weak, yet she found in her study upon control that, prior to the freeing of "attention," the selection of the correct path was largely dependent upon visual control and, at any time, under conditions of momentary distraction, visual perception played a major rôle. From Hunter[7] we have the very pertinent statement that "running the spatial maze . . . must require cues which have space localization as well as temporal location. In other words the rat must recognize in *terms of space* where he is in the maze. These spatially located cues may be anything but a kinæsthetic complex." A very good illustration of this same matter is to be found in the case of the rat (cited by Watson) which, having been once frightened at a particular place in the maze, continued to manifest emotional excitement at the same place long after the source of danger was removed. Here, without doubt, the animal's action was initiated by recognitory processes of a spatial character. Kuo,[8] using a form of multiple-choice, found that rats eliminated the incorrect path in a manner scarcely explicable without visual perception. To the rat a long alley was evidently a more circuitous path than a shorter alley. Most interesting of all to us is Kuo's statement that the "rat might *suddenly* cease going to the long path compartment." Thus Kuo indicates that there was an obvious extension of perceptual meaning, a visual apprehension on the part of the rat that carried it away from the longer pathway and caused it accurately to select the shorter. We are unable to decide what Peterson would have us accept as the significance of visual perception in the performances of his animals. At one place we find that "the rats evidently did not have time, nor adequate sense organs and conscious memories as a person would have to recognize and take note of external stimuli but resembled automatic machines in the quickness and uni-

[6] S. Vincent, The white rat and the maze problem, *J. of animal Beh.*, 1915, *5*, 24.

[7] W. S. Hunter, The temporal maze and kinæsthetic sensory processes in the white rat, *Psychobiol.*, 1920, *2*, 17.

[8] Z. Y. Kuo, The nature of unsuccessful acts, *J. of Comp. Psychol.*, 1922, *2*, 11.

formity of their responses." [9] How far Peterson is from the functional facts may be appreciated by a few months' observation of the use which a rat makes of vision in its orientation to physical objects, including its mates. Even an attempt to seize a rat will suggest the animal's use of its visual receptors. And so far as the matter of time necessary for recognition is concerned, we know that, at least for the human being, the meaning or significance is written across the object. Peterson a little later appears to relent, for he finds that there is a "tendency of animals with experience in mazes to proceed with less whole-souled response into *cul-de-sacs*. Let us suppose that as an animal enters a *cul-de-sac* it also receives certain stimuli of various kinds from the true path by which it departs. These stimuli may produce a weak partial response, or tendency to response, which does not immediately fade away. If this tendency persists until the rat emerges from the *cul-de-sac* it will, of course, enhance the impulse to take the true path and thus increase the probability of continued forward movement." [10] Peterson gives us no clear idea as to the nature of the "stimuli of various kinds"; neither does he give a clear-cut statement of the inferred tendency toward hold-over. One wonders what stimuli, if not visual; what function, if not perceptual, operated here. Again Peterson, forgetting the machine-like nature of his animals' performances, admits that possibly they also learn with training to utilize such factors as vague visual stimuli of the closed end of the *cul-de-sac*, eliminating thus, on the basis of vision, the short alleys earlier than the long.

Hunter [11] reports in another connection that certain of his rats, which had been placed for a short time each day in a box following their use and immediately prior to feeding, went, when permitted, to the box for eight days, because "the constant use of the box as a *link* in the food getting series . . . made it attractive in itself, or to state the matter differently,

[9] J. Peterson, Effect of length of blind alleys on maze learning, *Beh. Monog.*, 1917, *3*, 27.

[10] *Ibid.*, 39.

[11] W. S. Hunter, Note on the behavior of the white rat, *J. of Animal Beh.*, 1912, *2*, 141.

the *very perception of the box has acquired motive power.*" Here we have exemplified in the performances of Hunter's subjects, as well as in the other cases cited above, what appears to be a very important aspect of animal life; namely, the close fusion of perceptual apprehension of a visual character with actional performance so that the perception is not independent of action. According to Bentley, "an extremely important moment in psychological development is derived from the intimate fusion of perception and action. This fusion is strikingly presented in puzzling situations in which the child or other animal seeks a deferred and difficult outcome." [12] Essentially the same idea is expressed by Carr when, considering the process of acquisition in animals under repetition, he remarks that it is a matter "of organizing the items of experience into larger functional units. For example, the dog established an associative connection between the sight of the cane and the act of grasping it at one end and the rat learned to associate the act of lifting the latch with its sensory experiences of this object." [13] What we have, of course, in the performance of the animal is not an association of sensations, selected through sustained trial-and-error and stamped in through subsequent satisfaction, but an integration of a functional sort and one that is fairly simple. Neglect of this phase of animal activity by Thorndike, for example, who has had considerable influence upon the development of comparative psychology, has not been wholly conducive to an intelligent attempt to psychologize the animal. The tendency has been to obscure what appears to the writer to be the true aim of the comparative laboratory; to secure an accurate description of the likenesses and differences in the functional capabilities among similar and dissimilar animals from unlike levels of development. Such a functional description has no bearing upon the matter of intelligence, with which Thorndike was so much concerned.[14] There is no question of stupidity or of inferiority or superiority of animals as compared with man.

[12] M. Bentley, *The field of psychology*, 1924, 498.

[13] H. A. Carr, *Psychology*, 1925, 102.

[14] For a general criticism of Thorndike's theory of learning see K. Koffka, *The growth of the mind*, 1925.

Instead of a linking of mental processes we have rather, so it appears, the total performance of the animal; *i.e.*, the perception of situational objects, the movements of the animal toward food, freedom, shelter and the like, and the attainment of the end, woven into a unitary organic function. Concerning this integration of functions we have the very clear statement that "apprehension and action . . . stand as the most primitive and simple of our operations. It is likely that at an earlier stage of animal life the two functions were coincident, that the first operations involving these double resources were perceptive actions. Then the organism would, by a single performance, exert its mechanism in the direction of a given adjustive end which was vaguely and dimly represented by an apprehended object or quality." [15] Thus objects around the animal become apprehended and invested with significance only insofar as they contribute to the goal or end-result. The entire performance then becomes a way to satisfaction; and the objects, turns, openings, and so on, are invested with the meaning of "this-wayness" or "that-wayness." By ascertaining the character of the rat's response to changes in significant objects by way of shape, size and position, we may further determine the limits of the animal's functional capabilities. Finally, since there is, as we know, a total functional gain through repetition on the part of the animal, we may describe and explain in a like manner the animal's readjustment to new conditions.

Tolman has recently urged the acceptance of purposive categories in animal learning. He believes that the recognition of the fact that purpose and cognition serve as determiners of animal learning is a decidedly *fruitful* way of regarding performance. According to him, purpose or goal-seeking, is manifested not only through persistence until the end-result is obtained but also through a "tendency to *select* within limits the shorter of two or more alternative ways." [16] Under the influence of the goal the rat 'cognizes' and adjusts itself

[15] Bentley, *op. cit.*, 203.

[16] E. C. Tolman, Purpose and cognition; the determiners of animal learning, *Psychol. Rev.*, 1925, *32*, 287.

to various parts of the maze, and "these object-adjustments to the maze structure have meaning only with reference to the task of *getting to the food and getting there as quickly, both spatially and temporally, as possible.*" As we understand Tolman, he is insisting that the performance of an animal acquainted with the maze is directive in character and that the adjustments are made in terms of the *end to be attained.* We heartily agree with him, since this is precisely what we find to be true with our animals. The open door, in our case, meant a way to the goal—the satisfaction of eating—and this meaning was wholly perceptual. It was, that is to say, an extension under visual apprehension of a formerly restricted meaning to include the open door as the shorter way to food.

Conclusion.—We have sought to show that rats, among other animals, employ visual apprehension in a novel manner; and that the visual resources are not negligible in adjustment. We find clear confirmatory evidence both in the investigations of others, even in those who tend to deny visual perception to the rat, and in our own study. Furthermore, we agree with the exponents of the doctrine of the *Gestalt* that there is in the animals a very close fusion of the actional and perceptual functions; an integration, in fact, so intimate that one might rightly call it a simple organic function. We find, moreover, both in the literature and in our own experimental study, evidence that the readjustment of an animal to environmental changes may be characterized by a suddenness not explainable in terms of 'frequency' and 'recency'. It means that the situation has come to reflect for the animal, under an extension of visual perception, the satisfaction of a need or the attainment of an end. According to this way of thinking, the organism is not a blind mechanism propelled by kinæsthetic thrusts, but, like ourselves, it is capable upon demand of rapid adjustments to environmental changes. Such a way of regarding the white rat will tend, as we believe, toward a more intelligent understanding and interpretation of its performances. At the same time, it will duly recognize the functional utility of the visual apparatus.

THE EFFECT UPON DARK ADAPTATION AND VISUAL PERIODICITY OF ATROPIN AND HOMATROPIN

BY ROLAND C. TRAVIS

State University of Iowa

There are many factors to be considered in an attempt to explain dark adaptation in the human eye and that phenomenon in vision which resembles the refractory phase in neural function. Such physiological factors as fluctuations in the iris, the formation and bleaching of visual purple, contraction and expansion of rods and cones, afferent nerve excitation and conduction, synaptic resistances in the periphery and the brain, a central associative mechanism, and efferent pathways, are difficult of approach through accurate experimentation; and the psychological factors, after-image, contrast, binocular rivalry, attention waves, mental-sets, *etc.*, which are involved in dark adaptation and refractory period, are still less easily brought under experimental control due to their unaccountable variability from moment to moment and from day to day.

This study is primarily concerned with the question as to how the fluctuations of the iris influence dark adaptation and the refractory period in the dark-adapted eye. The experimental attack has centered itself on dilating the pupil by atropin and homatropin, thus ruling out the iris as a variable factor.

Troland[1] has found that the contractions and expansions of the pupil are a primary cause of the conscious fluctuation of stimuli of the same intensity and that there is a tendency for the pupil to adjust itself so that the retina receives a constant illumination under varying intensities. He asserts, however, that the fluctuation of the pupil fails to be effective in very bright and very dim lights, for which the pupil has attained, approximately, its maximal or minimal opening. Our results are in harmony with this conclusion for very dim lights.

[1] L. T. Troland, The theory and practice of the artificial pupil, *Psychol. Rev.*, 1915, 22, 167–176; The colors produced by equilibrium photopic adaptation, *J. of Exper. Psychol.*, 1921, 4, 344–390.

Reeves [2] found that it requires the pupil 5 minutes to open completely when the intensity is changed from bright sunlight falling on white paper to total darkness, and that it requires only 5 seconds for the pupil to close fully when the order of the same stimuli is reversed. The biological significance of the rapid closing of the pupil in bright lights as a means of protection for the retina is apparent; but the explanation of the longer reaction in very dim lights is not so clear.

In an analysis of Piper's work "Über Dunkeladaptation," Hecht [3] finds that the changes in the sensitivity of the retina in the dark are remarkably constant, not only in a given individual, but in different individuals as well. Age does not change the regularity of adaptation. The effect of drugs is practically negligible. Even persons with deficiencies in color-vision present a normal type of dark adaptation. Analysis of these and other data brings out the fact that little or no adaptation takes place after 45 minutes in the dark. In a later study [4] on foveal dark adaptation Hecht found that little or no adaptation takes place after 10 minutes in the dark.

It was found by Hess [5] that the receptor of the light reflex extends about 3 mm beyond the region of clearest vision in the retina, and that the cones act as the receptors for this reflex. The effector consists of two groups of muscles, one of which radiates like the spokes of a wheel and in contracting dilates the pupil; while the other group consists of a series of concentric bands which upon contracting narrow the pupil.[6] Illumination of the iris itself seems to have no effect on the light reflex, for when the optic nerves are destroyed by tumorous growth or trauma, light produces no change in the size of the pupil. There is some question as to the localization of the receptor for the light reflex on the retina, since it is an every-day occurrence that when the source of light is impinging on the eye from a large angle the pupil will rapidly contract. At such an angle the light must be falling on the peripheral retina. This does not eliminate the possibility, however, that the refracting media of the eye cause the light to fall on the cone-area as well as upon the periphery. A further fact to be noted here is that the stimulation of one eye effects the iris of both eyes, but more strongly the iris of the side stimulated.

Bourne [7] found that there is a definite causal relationship between the carbon dioxid content of the blood and the size of the pupil; the greater the CO_2 content in the blood the smaller the pupil and the smaller the CO_2 content the larger the pupil.

The fact that, under certain conditions, it requires the pupil 5 minutes entirely to open and only 5 seconds to close would lead one to believe that even with liminal stimuli the pupil may fluctuate in such a manner as to cause a refractory period of the visual mechanism. That is to say, more rapid closure may account, in part at least, for the time required between stimuli of like strength for the second to cause a visual response. So we may ask what part the fluctuation of the pupil plays in dark adaptation and in the visual phenomenon similar to refractory phase?

[2] P. Reeves, The rate of pupillary contraction and dilation, *Psychol. Rev.*, 1918, 25, 330–340.

[3] S. Hecht, Dark adaptation of the human eye, *J. Gen. Physiol.*, 1920, 2, 499–517.

[4] S. Hecht. Nature of foveal dark adaptation, *ibid.*, 1921, 4, 113–139.

[5] Hess, Untersuchungen über die Ausdehnung des pupillomotorisch wirksamen Bezirkes der Netzhant, *Arch. f. Augenhk.*, 1917, 48.

[6] T. V. Moore, *Dynamic psychology*, 1924, 58–61.

[7] W. Bourne, The relation between the width of the pupil and the carbon dioxid content of the blood, *Can. M. A. J.*, 1924, 14, 411. (Reviewed in *Intern. Sur. Ophthal.*, June, 1924.)

If the same relation between the time of pupillar opening and closing holds true in dim lights with dark adaptation as in the light-adapted eye, then the giving of a stimulus of low intensity would cause the pupil rapidly to contract, and some time must be allowed for the pupil to become adapted again in order to stimulate the retina a second time. This would not warrant the statement that the extent to which the pupil contracts varies with the intensity of the stimulus, but it would be in harmony with the statement that the contraction of the iris is a function, not of the absolute intensity of the stimulus, but of the state of adaptation of the eye. A stimulus that may not even effect the pupil or iris to the slightest degree in light adaptation, may cause a rapid contraction of the pupil when the eye is dark-adapted. If the statement is true that the extent to which the pupil contracts varies with the intensity of the stimulus then a very faint light would cause the pupil to contract relatively very little, and a relatively shorter time must elapse for the pupil to open again. This is out of harmony with the fact brought out by the previous investigation, as an explanation of refractory period, that the nearer the threshold the intensity of the stimulus becomes the longer the refractory period. The present facts seem to indicate that other factors aside from the function of the iris are operative in the phenomenon of refractory period, and also that the iris is only one of at least three structures subject to radical changes in dark-adaptation.

Method of Experimentation

The apparatus used in this study has been described in a previous investigation.[8] The apparatus is composed of a constant-current light, with absorbing screens interposed between the light and the observer. The intensity is further regulated by an iris diaphragm the opening and closing of which are controlled by a pointer on a scale with gradations from 0 to 28. When the diaphragm is completely open fifty times as much light intensity is falling on the end screen as

[8] R. C. Travis, A phenomenon in vision similar to refractory phase, *Psychol. Monog.* (*Univ. of Iowa Studies in Psychology*, 10) 1926.

when it is completely closed. A small fixation-light is used to hold the observer's attention in the direction of the stimulus. The stimulus proper has a much larger area, and is turned on periodically. It appears to the observer as a round patch of white light. A circuit and key are used for the observer's response. The entire apparatus is in a light-proof room.

Observations on two subjects make up the experimental data in this study. Four experiments were made on each observer, two while the observer's eyes were under the influence of atropin (in the case of A) and homatropin (in the case of B), and two when the pupils were again normal.

The first experiment began as soon as the observer entered the dark-room and continued about 40–50 min. This experiment attempted to show the relation between the intensity of the stimulus and the amount of dark-adaptation with a constant interval of 30 sec between stimuli. In other words, the sensitivity of the visual mechanism was determined at successive stages of dark-adaptation. Each subject observed in this experiment twice, once while his pupils were completely dilated, and again when they were normal.

In the second experiment it was attempted to show the relation between the intensity of the stimulus and the refractory period, with dark-adaptation relatively constant. This was done immediately after the first experiment while the observer's eyes were dark-adapted. A period of ten minutes was allowed to rest the observer's eyes before the second experiment began. The second experiment continued not longer than 12 minutes.

In each experiment the fixation-light was on continuously, and the observer looked steadily at this point with his head resting easily in the headrest. Each time the observer saw the stimulus (which appeared every 30 sec, without the observer's knowledge of the time-interval), he responded by pressing the key. The observer was credited with this response when he responded two times to that intensity and failed to respond to a lower intensity. This was taken as the threshold for that period of dark-adaptation. The threshold was accurately determined every 2 to 6 minutes during the course of dark-adaptation.

The observer's threshold in the second experiment was obtained with a time-interval of 120 sec between stimuli. This interval was empirically determined as the longest required time-interval. As a matter of fact, the time-interval at the threshold was 100 and 110 sec for *A*, and 65 and 95 sec for *B*. The refractory period was then determined for intensities above the threshold up to 28, the upper limit on the scale.

The results obtained from a third observer *C* are included for comparison. *C* observed in each experiment four times on four different days. The same conditions of the experiment were maintained for *C* as for observers *A* and *B*, except that *C*'s pupils were normal throughout.

Interpretation of Results

Observer *A*'s pupils had received the last dose of atropin for complete dilatation one hour before the experiment began, and previous to the experiment had been exposed to the bright sunlight. Curve (1) in Fig. 1 shows the relation between

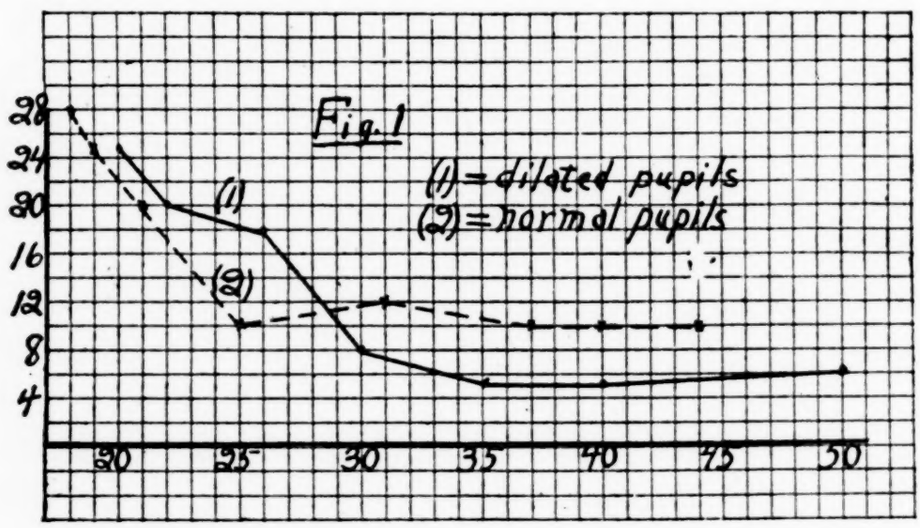

Fig. 1

intensity and dark-adaptation with a constant interval of 30 sec between stimuli, for the dilated pupils. The ordinates are the units of intensity (scale on iris diaphragm), the abscissae the dark-adaptation in minutes. Curve (2) shows the same relation for the normal pupils 17 days after dilatation,[9] where

[9] The effect of atropin begins to diminish in about 4 days, and complete recovery is obtained in about 15 days.

the observer was exposed to the bright sunlight previous to the experiment.

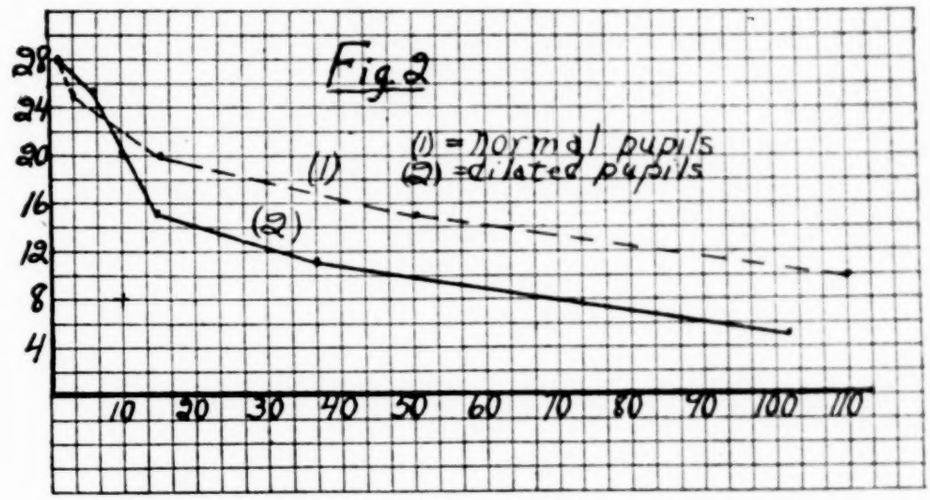

FIG. 2

The two curves cross about the center, but their general slope is approximately the same. The threshold is about five points lower for the dilated pupils as compared with the normal. This difference has little significance when compared with the variability in the relation between intensity and dark-adaptation in the case of observer *C* (Fig. 5), and the threshold varies from 1 to 10 on the same scale (Fig. 7). The curves cross each other at several points (Fig. 5), but they all have about the same general slope.

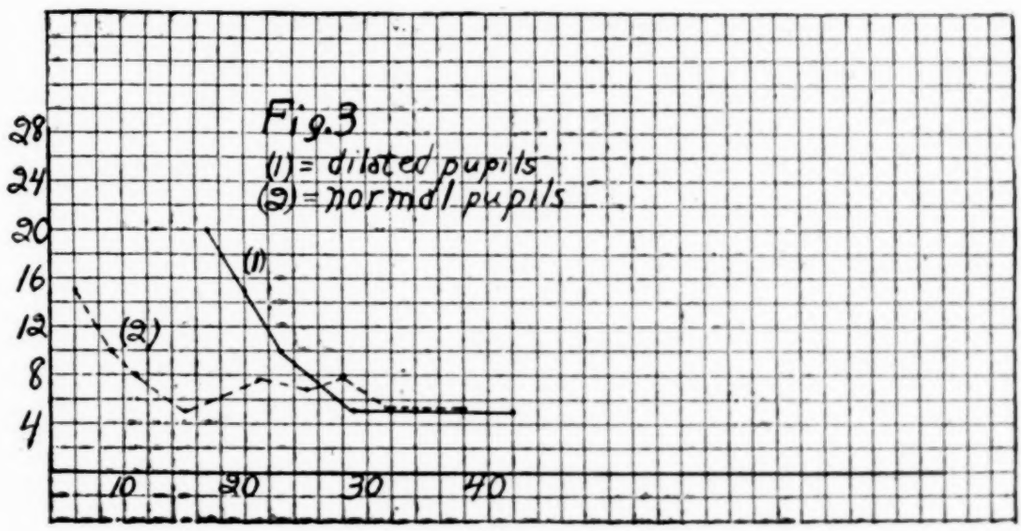

FIG. 3

Observer *B*'s pupils received the last dose of homatropin for complete dilatation 3 hours before the experiment began, and previous to the experiment had been exposed to the bright sunlight. Curve (1) in Fig. 3 shows the relation be-

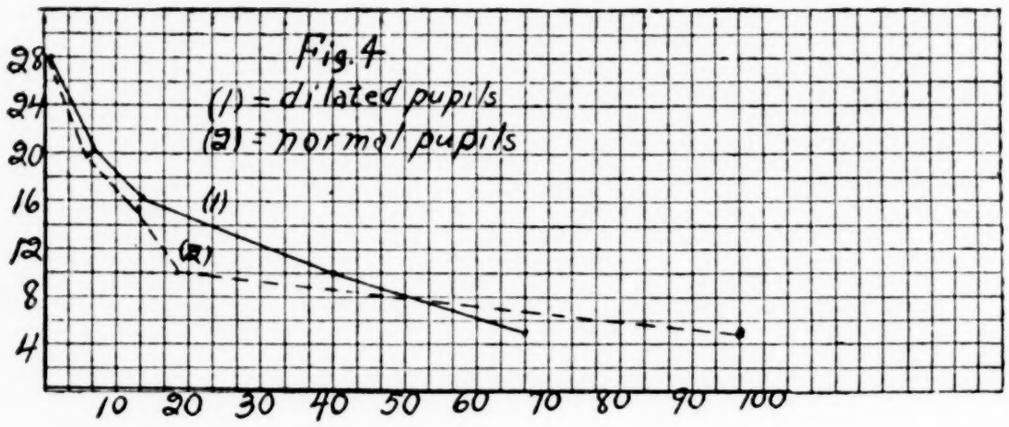

FIG. 4

tween the intensity and dark-adaptation. Curve (2) shows the same relation for the normal pupils 10 days after dilatation,[10] where the observer was previously exposed to the in-

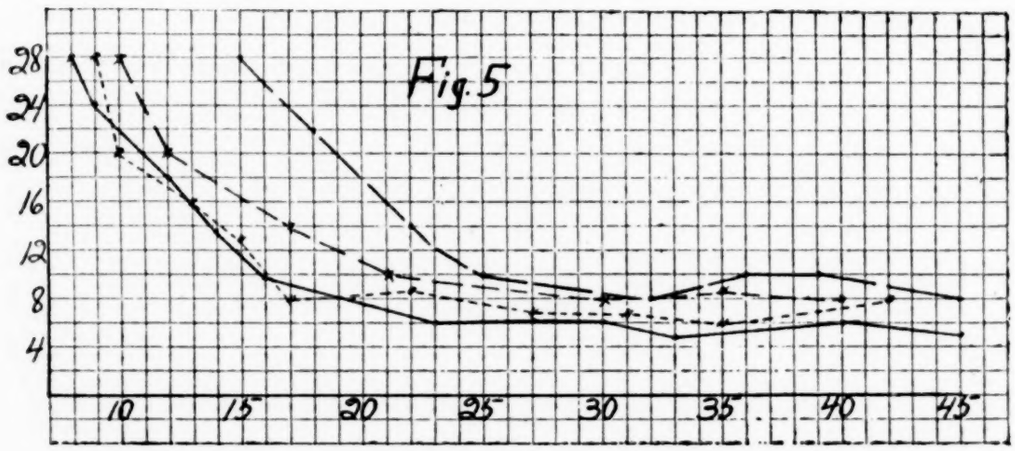

FIG. 5

tensity of a cloudy evening. The general slope of the curves is the same and the thresholds for both trials are the same (5 units). It will be seen, however, that, for the dilated pupils, dark-adaptation went on for 17 minutes before stimulus-

[10] The effect of homatropin begins to diminish in about 12 hours and complete recovery is obtained in about 2 days.

intensity = 20 brought out a response; whereas, in the normal pupils, the intensity = 15 was responded to after 6 minutes in the dark. This is readily accounted for by the fact that *B* was exposed to bright sunlight previous to the experiment when his pupils were dilated, and it was late in the afternoon on a cloudy day when he was tested with normal pupils. The same situation occurred in the case of *C*. The three curves in close proximity (Fig. 5) represent results of tests made in the evening, the other curve represents previous exposure to the sunlight.

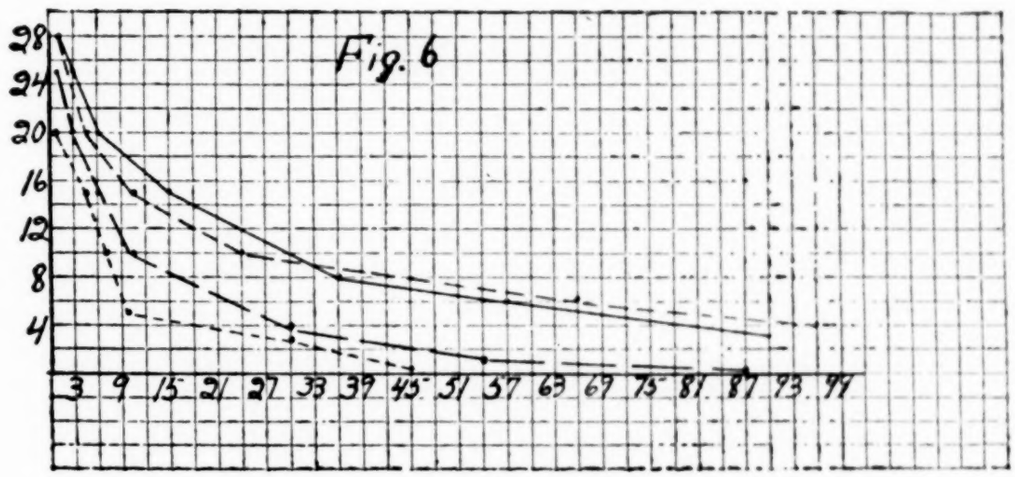

FIG. 6

The conclusion seems to be warranted that the dilatation of the pupils by atropin and homatropin has little or no effect on the dark-adaptation of the eye, other than making the process a little more prolonged.

Figures 2 and 4 bring out the relation between the intensity and the refractory period, with dark-adaptation relatively constant (dilated and normal pupils of observers *A* and *B*). Curve (2) in Fig. 2 and Curve (1) in Fig. 4 show the relation between the intensity and the refractory period for the dilated pupils. Curve (1) in Fig. 2, and Curve (2) in Fig. 4 show the same relation for the normal pupils.

There is a small but significant difference between the curves for dilated and those for normal pupils. In observer *A*, the refractory period is a little less throughout for dilated pupils than it is for the normal, in spite of the lower threshold

for the dilated pupils. In *B*, the refractory period is about one-third less at the threshold for the dilated pupils, yet at intensity = 10 the refractory period for the normal pupils is about one-third that of the dilated. This same sort of variability is shown in Fig. 6 on the curves obtained from observer *C*.

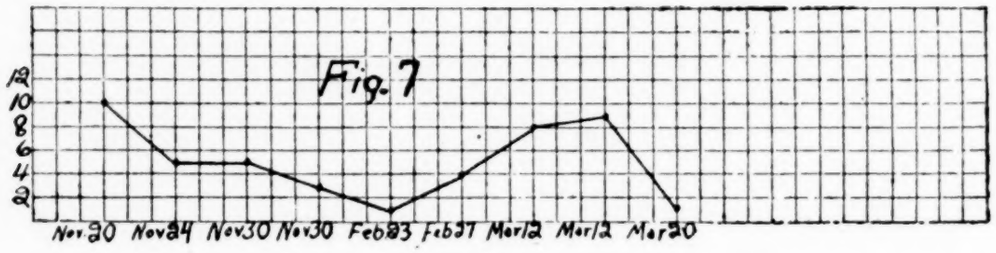

FIG. 7

The greater part of the difference noted between the normal and dilated pupils in regard to dark-adaptation and refractory period can be explained as normal variability in the function of dark-adaptation and of the refractory period. This variability in function is no doubt due to such changing physiological conditions as fatigue, metabolic rhythms, and neural and muscular periodicities. The fluctuations in the iris and in the visual threshold and the refractory period of the scotopic eye are phenomena which future experimentation will show to be attributable to more basic fundamental periodicities comparable to the refractory phase in neurone and muscle and to the maintenance of chemical balance in all tissues and organs.

SUMMARY

This study reports experiments on the effect upon dark-adaptation of the eye and the visual phenomenon similar to refractory phase, as produced by dilatation of the pupil by atropin and homatropin. Four experiments were made on two observers; two for the dilated pupils and two for the normal. The first experiment had to do with the effect of dilated pupils upon dark-adaptation; the second with the effect of dilated pupils on the refractory period.

Most of the difference noted between the results obtained on the dilated and the normal pupils can be explained on the

normal variability in the function of dark-adaptation and the refractory period, such variation being probably due to changing physiological conditions.

The dilatation of the pupils by atropin and homatropin seems to have little or no effect upon dark-adaptation except to make it a little more prolonged. This prolongation is due to the over-stimulation of the retina by light caused by a too-long sustained dilated pupil. It also seems safe to infer that the iris plays a minor part, if any, in the visual phenomenon which resembles refractory phase. Because of the limited number of cases involved the conclusions drawn from this study must be taken as suggestive and not as final evidence.